

CA20NEC 12

-75P03



3 1761 11892478 6

Working Paper Series

A RESEARCH AGENDA
IN
HEALTH CARE ECONOMICS

R. D. Fraser

October, 1975



Ontario Economic Council

Toronto, Ontario



CA20NEC 12 13
-75P04

Government
Publications


A RESEARCH AGENDA
IN
HEALTH CARE ECONOMICS

R. D. Fraser

October, 1975

Working Paper No. 3/75

The opinions expressed in this paper are the author's
and do not necessarily reflect the views of the Ontario
Economic Council.



Digitized by the Internet Archive
in 2023 with funding from
University of Toronto

<https://archive.org/details/31761118924786>

TABLE OF CONTENTS

- Chapter I: Introduction
- Chapter II: The Behavior of Health Care Personnel
- Chapter III: Resource Allocation in the Hospital Sector
- Chapter IV: Consumer Behavior in the Health Care Sector
- Chapter V: The Behavior of Government
- Chapter VI: The Organization of Non-Hospital Ambulatory Care:
Solo versus Group Practice
- Chapter VII: The Substitutability of Human Health Care Resources
- Chapter VIII: Advances in Health Care Science and Technology
- Chapter IX: The Distributional Impact of Government Funded
Health Care
- Chapter X: Conclusions

Chapter I

INTRODUCTION *

I. The Delivery of Health Care: An Economic Problem

Whether because of rising levels of income and education, a changing age distribution of the population, changing social attitudes, or perhaps because of the establishment of government as the principal funder of health care, expenditures on health care have been rising rapidly in Canada in the last decade and a half. As shown in Table I-1, total expenditures on health care in Ontario have risen some 263 percent over the period from 1960 to 1971. Though there are differences in the extent of the increases among the different components of the sector, each is seen to have experienced fairly rapid growth; the hospital area, however, is clearly shown to have had the most rapid growth in expenditures. We might also note that the rapid increase in health care expenditures over the past decade has left Canada in the position, which it has occupied for sometime, of being the country devoting the largest percentage of its resources to the provision of health care.

* I wish to acknowledge the many helpful comments made by Kerry Adams and Michael Prime on the first draft of this manuscript; the several fruitful discussions on health research priorities held with H. Walker of McMaster University; the comments made by a wide variety of individuals brought together for the seminar arranged by the O.E.C. on August 27, 1974; and the comments made in the session on research priorities of the Symposium on Health Care Economics held on September 5 and 6, 1974 at Queen's University in Kingston.

Table I-1: Total Expenditures on Personal Health Care: Ontario and Canada, 1971

Expenditure Category	Ontario			Canada		
	Expenditure \$000	Per Cent of Total	Per Cent Increase 1960-71	Expenditure \$000	Per Cent of Total	Per Cent Increase
(1)	(2)	(3)	(4)	(5)	(6)	(7)
All Hospitals	1,189	59.5	282.3	3,152	61.7	273.0
Physicians' Services	508	25.4	257.7	1,236	24.2	248.2
Dentists' Services	142	7.1	189.8	299	5.9	174.3
Prescribed Drugs	160	8.0	233.3	422	8.3	219.7
Total	2,000	100.0	263.0	5,110	100.0	254.4

Source: Government of Ontario, Report of the Health Planning Task Force (Toronto:1974), p.53.

Not only have these expenditures been rising rapidly but also they have been rising more rapidly than the level of economic activity in the economy at large. As shown in Table I-2 some 6.0 percent of GNP was allocated to the provision of health care in 1961 and some 7.3 percent in 1969, the latest year for which comparable data have been compiled. Thus without question, the health sector may be labeled not only one of the principal sectors of activity in the Canadian economy, but also one of the faster growing sectors.¹

The sheer size and rapid growth of this sector, coupled with the role government now plays in it as the principal funder of expenditures on health care, flag it as a potentially fruitful area for research concerning the efficiency with which the resources of the sector are organized and employed. The fundamental question that requires answering is: "Are health care resources yielding benefits commensurate with expenditures on them?" A recently completed international study² of the relationship between infant mortality (as one measure of health status and thus the output of the health sector)

1. Furthermore, this recent rate of growth seems unlikely to abate in the near future. Expenditures on health by the government in Ontario are expected to increase from 2.3 billion in 1974 to some 3 billion in 1975, or an increase of some 30 percent as reported in "Ontario Removes Ceilings on Budgets of Hospitals", Globe and Mail, Tuesday, October 29, 1974, p.1.

2. R. D. Fraser, "An International Study of Health and General Systems of Financing Health Care," International Journal of Health Services, Vol. 3, No. 3 (1973), pp. 369-97.

Table I-2.- Total Expenditures for Health Services as a Percentage of the Gross National Product: Seven Countries, Selected Periods, 1961 - 69.

Country (Ranked by Percent of GNP)	WHO Estimates ^a		SSA Estimates ^b		Average Annual Rate of Increase in Health Expenditures ^c	
	Year	Percent of GNP	Year	Percent of GNP	Year	Percent Rate
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Canada	1961	6.0	1969	7.3 ^d	1961-69	13.2
United States	1961-62	5.8	1969	6.8	1962-69	10.1
Sweden	1962	5.4	1969	6.7	1962-69	14.0
Netherlands	1963	4.8	1969	5.9	1963-69	16.1
Federal Republic of Germany	1961	4.5	1969	5.7	1961-69	10.3
France	1963	4.4	1969	5.7	1963-69	14.9
United Kingdom	1961-62	4.2	1969	4.8	1962-69	9.5

a) Brian Abel-Smith, An International Study of Health Expenditure. WHO Public Paper No. 32 (Geneva: 1967).

b) Joseph G. Simanis, "Medical Care Expenditures in Seven Countries", Social Security Bulletin, (March 1973), in which the updating of the U.S. Department of Health Education and Welfare, Social Security Administration is reported.

c) Office of Health Economics, London, "International Health Expenditure 11", Information Sheet No. 22, (May, 1973), Table 2.

- d) This figure is somewhat higher than it should be according to a recent publication of the Department of National Health and Welfare, National Health Expenditures in Canada, 1960 - 1971 (Ottawa: 1973), Table A4, p.28. In it the figures for 1961 and 1969 according to a different definition of health expenditures than that used in Table 2, above are 5.7 and 6.4 percent respectively. On the basis of the percentage increase from 1961 to 1969 thereby indicated, and assuming that the 1961 figure in Table 2 is correct, the comparable figure for 1969 would be 6.7 percent.

Source: Government of Ontario, Report of the Health Planning Task Force (Toronto, 1974), p.52.

and measures of the input of health resources¹ indicated that Canada ranked 23rd of the 25 developed countries studied with respect to the rate of infant deaths per 1000 live births. Though the results of this study do not themselves offer compelling proof of inefficiencies in the organization of the existing bundle of resources in the health sector, they are supportive of the view that this is a principal area for further research.

Equally important, these results and others indicate that this whole field of health care is an important one for economic research. In addition to the essentially medical aspects, there are fundamental, economic dimensions to the problems involved in delivering health care. Notwithstanding the recent comments of the Federal Minister of Health, the wants of the Canadian population for health care can be described proximately as insatiable. Resources for satisfying these wants are, without question, scarce. Furthermore, there are alternative ways of combining the physical and human resources of the health sector to provide health care in response to a given health problem.

Clearly then, the delivery of health care can be described as an economic problem which requires answers to the following broad questions commonly posed and analyzed by the economist:

1. Which health goods and services (including the level of quality), and how much of each, should be provided?

¹. Also included in the analysis were factors to control for social, economic and demographic determinants of health status.

2. How should they be provided?
3. For whom should they be provided?

In the private economy, competitive forces working largely through factor and product markets provide an answer to these questions and generally ensure reasonably efficient provision of the goods and services in question. Within the market system, prices play the fundamental role in allocating resources, both those used by producers and also those purchased as final goods and services by consumers. The distributive aspect namely the "for whom" question, is also dealt with by the market system but not in an unmitigated form. The results of market processes are attenuated by public policy.

The following kinds of choices, which would be made in the private economy by the decentralized decision making of suppliers and consumers with the essential assistance of prices, are the principal types of choices that must be made with respect to the health care sector:

1. the choices of physicians (and other providers of health care) of the stream of services to recommend and/or provide to the patient; choices from the given range of alternative health facilities, manpower and treatment programmes;
2. the choices of consumers as to the stream of services they "request" for the amelioration of a given health problem; choices from a given array of health facilities and manpower. Both under-utilization (of, for example, preventive

health measures) and over-utilization (including the use of more highly qualified manpower than is justified) can result if mechanisms do not exist for modifying the unrestrained choices of consumers;

3. the choices with respect to new types of, or rearrangements of, existing health care facilities and the expansion of their availability. This area of choice includes decisions on the number and location of coronary care units, renal dialysis units, etc. as well as the choice of the number and mix of hospitals, community health clinics and similar facilities;
 4. the choices with respect to new types of health care manpower, or with respect to expanded roles for existing health occupations, and with respect to the existing levels of availability. This area of choice involved consideration of the adequacy of educational programmes, potential manpower shortages, etc.;
 5. the choices of new health goods and services that might become part of the range of service streams open to the physician or other health professional to prescribe and/or, provide;
 6. the choices of systems and levels of remuneration of the health professions;
 7. the choices with respect to the geographic, demographic, cultural, and socioeconomic, distribution of health care.
- This is a question not only of the availability of health

facilities and manpower but also of the impediments, cultural and otherwise, to their use by certain segments of the population;

8. the choices with respect to the overall size of the health sector.

When an industry becomes part of the public economy the economic problem of allocating resources is not thereby solved or eliminated. Mechanisms are still required to answer the economic questions set forth above. In particular, a general system is required for guiding the various economic agents in making choices of the kinds just set forth.

It is in this framework then that an attempt is made in the ensuing paper to set forth, from an economic perspective, an overview of several selected areas of research that are thought to be of prime importance. In particular, the areas examined are those in which the tools of the economist are thought to have special relevance and usefulness.

Before setting out to discuss six areas of health care research which we argue are of direct relevance for economists, we present an overview of the health care sector. For this purpose we employ the taxonomic framework frequently used by economists working in the general area of industrial organization, namely that of Market Structure, Market Conduct and Market Performance. In so doing, we are able to emphasize most clearly the relative lack of knowledge about the productive

processes and interrelationships among the principal economic agents of the health care sector.

This method of describing the health care sector is clearly experimental. Can the taxonomic framework of the industrial organization economist be used with advantage to describe the nature of the health care sector? At this time, I strongly believe that it can serve as a useful framework for gaining a comprehensive overview of the sector. It is taxonomic, however, and as such differs from alternative frameworks that might have been used. For example, one might have arrayed the major policy problems of today's health sector and therefrom delineated research problems. This tack would, however, be largely short run in nature especially in view of the nature of the research process as opposed to operational management.

At the same time, we have adopted yet a third alternative framework in the more detailed discussion of research priorities that is contained in Chapters II through X. Therein we use a framework that falls first from the two broad goals of allocative efficiency and distributional equity and secondly subdivides the issues of allocative efficiency into issues that relate directly to the major economic actors of the health care sector.

II. An Overview of Market Structure, Conduct, and Performance in the Health Care Industry.

Introduction

The delivery of health care must be counted as one of the principal sectors of economic activity in our economy if for no other reason than its sheer size. We have thus employed the "market structure-conduct-performance" framework to describe this sector. In this framework the elements of the structure of an industry are related to the behaviour or conduct of its principal economic units and in turn both the structural and behavioural elements are related to the performance of the industry.

The basic elements of this framework are, however, those describing market structure. In describing it one is concerned with determining how close a particular industry is to being "competitive". The condition of many buyers and sellers none of whom can influence market price is examined by estimating the number and size distribution of buyers and sellers as (principally) indicated by measures of the concentration of output and consumption. The condition of factor mobility is considered in terms of the nature of the barriers to entry to the industry, as seen for example in the presence of significant economies of scale or the presence of government legislated barriers to entry. The condition of product

homogeneity is considered in terms of the nature of the product produced and in particular the extent of product differentiation, through advertising for example. The fourth principal condition of perfect competition, namely that of perfect knowledge is considered more generally in terms of the complexity of the product produced and the existence of means of informing the consumer about the nature of the product in question. Finally, underlying preconditions affecting the "efficiency" characteristics of perfect competition such as the assumed independence of utility and production functions (not only as among suppliers and consumers taken separately, but also as among these two groups), are also taken into account.

In setting out this brief descriptive framework of the health care industry we will not be noting the many detailed sources of reference to the statements made. To be done properly this would require by itself a full-scale industry type study of the health care sector. On the other hand, this descriptive model does offer significant potential for assisting in a general understanding of the economic processes of the health care sector and the basic issues surrounding them. We will point out the numerous gaps in our knowledge of what are, or at least what are presumed to be, its principal elements.

The discussion will be framed in rather narrow economic terms not because economic considerations are, or should be, ultimate or exclusive criteria but because we wish to emphasize those aspects concerning which an economic approach enjoys a comparative advantage. Given the lack

of qualification and brief compass, our remarks will be much stronger than is warranted in the light of our present ignorance. But this is advantageous because our fundamental thesis is that not much, and certainly not enough, is known about the economic behavior of the health care sector. Sharp and controversial conclusions for which there is some evidence but with which the reader may well disagree will help to indicate both the dimensions of our ignorance and the sort of knowledge that must be gathered to generate firmly based judgments. How much can be learned to advantage has, itself, an economic dimension. In our judgment, much that is useful can be known.

Market Structure

A.1. Concentration of Buyers

The consumers of health care must be judged to be fairly widely distributed throughout the population. Surveys of health status point to the importance of demographic, social and economic factors as determinants of the utilization of health care. Much more information and analysis is required and could be accomplished. With respect to the basic proposition, however, it is clear that there has generally been little concentration of buying power in any of the principal components of the health care sector. The one possible exception might be the purchase of pharmaceutical products and other equipment

and supplies for use by hospitals.

With the introduction of hospital insurance in the early 1960's and medical care insurance in the latter part of that decade, the nominal concentration of buying power would seem to have been completely reversed with government, now installed as the principal funder of health care, potentially a monopsonist. Just how effective this new concentration of buying power has been is open to question. It can be argued, as for example in the case of Ontario, that government has acted in the passive role of an insurer and has used but a negligible part of the buying power that it would seem to possess. The pros and cons of whether this power should be used are of course distinct from the question of whether it has in fact been used.

A.2. Concentration of Sellers

In marked contrast to the buyers, the providers of health care can be said to be highly concentrated. With respect to hospital care, the population of many communities effectively has access to only one or at most a few hospitals. Furthermore, the presence of associations of hospitals suggests a less than independently determined supply of hospital services as between different hospitals. But, one cannot be certain of the degree of interdependency.

In the non-hospital, ambulatory care sector, although there are many physicians, they cannot be said to act independ-

ently. The presence of approved fee-for-service schedules attests to the success of the Medical Associations and Colleges of Physicians and Surgeons in formulating collusive selling practices. The same can be said of the many sub-groups of physicians and also for the dental care component of the health care sector. The quantitative importance of such practices awaits further study, however.

With respect to the supply of pharmaceutical products, the concentration of sellers is high. This component of the health care sector can most accurately be described as oligopolistic rather than competitive.

B.1. Barriers-to-Entry: Economies of Scale

The nature and extent of economies of scale in the provision of health care is subject to much debate and requires further research for almost every component of the sector. We simply do not know very much about these matters. The possible exception is the production and distribution of pharmaceutical products. In this case there have been several studies done but as yet there is no general agreement that the production processes in this area are indeed characterized by significant economies of scale and that these serve as an important barrier to entry.

In the provision of hospital care, non-hospital ambulatory care, and dental care, the evidence thus far accumulated offers far from compelling proof that economies

of scale exist in a significant way. The weight of the evidence in all three areas suggests that modest economies of scale are present. The "weight" of the evidence is not, however, weighty. This is in contrast to the weightiness of intuitive, a priori arguments, most of which would lead one to expect to find significant economies of scale in all these processes, at least over initial ranges of production.

Once again, we simply do not know with any great certainty what the facts of the case are. The assertion and arguments about the greater efficiency of group practice turn very much on not only the existence of potential economies of scale but also on the probability of these being exploited.

One aspect of the nature of economies of scale that recent analysis has begun to shed light on is that related to the human capital involved in the provision of certain complex surgical operations. For example, unless a certain number of heart or kidney transplants can be done each year by a given surgical team, the productivity of the team as seen principally in the quality of care provided is said to fall significantly. This would appear to be an area of potentially high dividends to much further research.

B.2. Barriers-to-Entry: Absolute Cost Advantages

Absolute cost advantage is often seen to play an important role in retarding the entry of new firms to a

monopolized industry in the economy at large. These cost advantages are sometimes seen in advantages related to the large scale advertising done by large and established firms (where such advertising is itself characterized by economies of scale), in advantages related to the establishment of preferred positions in the capital market, and in advantages related to the exclusive control of a particular factor of production where such control might be in the form of the ownership of the principal source of a scarce raw material.

In the health care sector, the pharmaceutical industry is probably the component in which these various cost advantages likely exist and thereby could retard the entry of new suppliers. In particular, the existence of "patents" is probably the most significant source of cost advantages and in turn barriers to entry. Once again, one cannot be certain about the quantitative effects of this.

Certain parts of the hospital care component may in the past also have been characterized by the existence of absolute cost advantages as seen in the use of nursing and other services as supplied by the members of different religious groups at wage rates below market rates.

B.3. Barriers-to-Entry: Government Legislation

By far the most overwhelming barriers to entry in the health care sector would seem to be those directly related to the presence of government legislation. In the hospital sector,

the establishment of new hospitals has in many provinces been directly regulated by government legislation.

In the non-hospital, ambulatory care sector, government legislation has given to the Colleges of Physicians and Surgeons the responsibility for setting out licencing procedures for physicians including the educational requirements and various regulations that accompany these procedures. Indeed, in general, manpower in the health care sector is circumscribed by a vast body of government legislated registration, certification and licencing procedures.

B.4. Barriers-to-Entry: Summary

Though some kinds of health care such as nutritional advice, family planning and counselling, mental health care, and so forth are obviously provided by economic agents outside of the traditionally defined health care sector, and their provision is thus not generally subject to legislated barriers to entry, by far the major components of health care are provided under conditions in which barriers to entry are present and are relatively high.

C. Product Homogeneity

Though the existence of government funding may appear to be promoting the provision of a more standardized package of health care to persons with given diagnoses, as for example

in the hospital sector, the very nature of health care as principally a service or set of services provided largely by human capital precludes the attainment of anything close to product homogeneity. Indeed, with respect to physicians of roughly the same age, with roughly the same educational and training background, and employing roughly the same mode of practice, the quality of care provided has been demonstrated to vary widely. The same is likely true of much of the entire spectrum of health care.

The pharmaceutical area is one on which we might elaborate. In it, there would seem to the casual observer to be a much greater potential for achieving product homogeneity because firms in this area are involved in producing well defined products. The oligopolistic nature of this industry and especially the role played by advertising, and product innovation and development seem however to attest to the success of pharmaceutical firms in differentiating their products from those of competing firms.

D. Information Flows

The condition of perfect knowledge is of course never satisfied with respect to any market of interest. The level of information about prices and the characteristics of competing products including especially their quality is nevertheless sometimes relatively high. This may be because the good in

question is a relatively simple one whose nature is easily comprehended by the vast majority of the population, or, because the good, though complex in nature, is one about whose characteristics satisfactory information can be and is made available. The mechanisms for providing this information may be found in the industry itself as seen for example in the role played by advertising, or may be found in an associated industry whose principal function is to supply the information that consumers wish to have about the product in question.

In the health care sector, the flow of information about the quality of the services rendered by various types of health care manpower in various health care facilities is characterized by its imperfect nature, if not its complete absence. At one stage in the working life of health care personnel, namely on qualifying for certification and a licence to practice, some minimum standards are of course assured. Generally, however, the value of this assurance depreciates rapidly. The one possible exception to this statement would be found in the pharmaceutical component. In it, advertising has clearly played a role in providing information about the quality of different products. In addition, government has been involved both in stating standards for the introduction of new drugs and more recently in delineating substitute drugs as in the case of the PARCOST programme in Ontario.

E.1. Externalities in Consumption

It has long been argued that there exist significant external economies in the consumption of health care that militate against its "efficient" provision through the mechanism of the (unregulated) market. The oft used example is the external effects of the medical care provided by a physician to those who have contacted a contagious disease. With inadequate treatment, the carrier of typhoid, the bubonic plague and so on could yield tremendous negative benefits to both other suppliers of health care and also the public at large.

E.2. Externalities in Production

In a roughly analogous way, it is often argued that a significant (though not necessarily large) part of the spectrum of health care goods and services is characterized by external diseconomies of production. The traditional example would be the case of the provision by one individual or group of individuals, of services designed to clear a malaria swamp. Other individuals, not in the group buying such services, will nevertheless be recipients of the better health associated with the eradication of the swamp. Again, the existence of these externalities is often used to justify the supplanting

of the market system by government action in the provision of health care.

Market Conduct

A.1. Pricing and Output Behaviour: Hospitals

Since almost all the hospitals in Canada are being funded principally out of government treasuries through the hospital insurance plans, pricing behaviour has almost become a relic of the past. The same is not true of the output decision. Both the quantity and quality of each of the wide range of potential hospital services must be determined for each hospital.

The growing but yet small body of literature describing this decision-making process in what are not only non-profit, but also largely non-revenue generating institutions has begun to include some common strands of thought. First, the objective function of the hospital, including its decisions on output, is seen to be some amalgam of the objective functions of at least three principal parties: the physicians who have hospital privileges in the institution, the hospital administrator, and the Board of Trustees or Governors. Secondly, the role played by the physician group mentioned above is likely to be the preponderant one. And thirdly, the objective function of the physician group, as it relates to output, will not be a simple one since this group is itself composed of

several distinguishable subgroups. For example, there are likely to be differences in outlook between the physicians who are essentially hospital based specialists and those who primarily practice general medicine outside of the hospital but who like to use the hospital facilities from time to time. Indeed, the granting of hospital privileges to potential competitors would seem to have the potential for placing existing medical staff into a position of conflict-of-interest which would be disadvantageous to society.

In spite of the expected and sometimes documented differences in the objective functions of the three principal parties noted above, there appear to be some common objectives with respect to output. For example, there is likely to be a consensus directed at increasing both the quantity and/or the quality of hospital services available without regard to whether the benefits to the consumers are sufficient to outweigh the costs of providing them. That its numbers of hospital beds of various kinds place Canada in nearly the top position of beds per 1000 population, is certainly evidence consistent with output maximization being an element in the objective function of a hospital.

A.2. Pricing and Output Behaviour: Non-hospital Medical Care

In the second most important component of health care sector, namely that of providing non-hospital ambulatory care, the "pricing" part of the pricing and output decision is very

much alive. Though government has now effectively replaced the patient as the financial supporter of these medical care services, the providers are still remunerated through the fee-for-service system of determining payments. Furthermore, the physicians as a group would seem to have a disproportionately large voice in the determination of the absolute level of the schedule and an almost unopposed voice in determining the relative levels of fees for various services.

With respect to the overall level of fees, it is possible, and obviously of crucial importance to any government funding medical care, to enquire whether physicians in general, or particular sub-groups, are currently facing leisure-income decisions that could be characterized by a backward bending supply curve of labour effort. That is, as physicians' incomes rise is there a tendency to abandon the life style of the harried professional by reducing the amount of time devoted to patient care? We don't know.

Certainly, overall levels of income for physicians are on average high enough to indicate a disequilibrium in this labour market. Moreover, there is a growing unmet demand for places at medical schools in Canada; there is an immense flow of immigrant physicians to Canada (and from the other provinces to Ontario). The income levels of physicians have been increasing more rapidly than both those for professional occupations in total and than those for the population as a whole.

With respect to relative pricing and output decisions, there seems to be strong support for the contention that certain sub-groups of services are disproportionately priced too high with resulting surpluses in the supply of manpower in these areas and in contrast that there are others priced too low. An example of the former would be some of the hospital based specialties in the surgical area and an example of the latter would perhaps be the area of general medical practice. Some of these problems can be related to demonstration models followed in medical education programmes; some to the introduction of a simple-minded health care insurance scheme that did not take account of the varying levels of charity and bad debts which characterized the different subgroups of physicians; and some to the advances of medical science especially in the hospital sector. All of these causes, however lead to problems only when coupled with a fairly rigid system of determining the relative fees in the fee-for-service schedules.

B. Advertising

The role of advertising is probably only significant in the pharmaceutical products component of the health care sector. In this area, it probably has much the same characteristics as advertising in other sectors of industrial, economic activity. It is therefore likely amenable to analysis using roughly the same methodology.

In the other components of the health care sector, advertising is either negligible or is prohibited completely. Indeed, whereas in the pharmaceutical products area one is faced with the normative question: "Is there too much advertising?"; in the other components one is faced with the question: "Is there too little advertising?". We have mentioned above that information flows are generally significant determinants of the efficiency with which markets operate, and further that advertising can play a significant role in facilitating the flow of information. That advertising is all but prohibited in the major components of the health care sector is thus cause for concern and should be ample cause for analysis of the impact of this prohibition on the efficiency of the delivery system.

C. Research and Innovation

As a principal behavioural trait, research and innovation probably plays a crucial competitive role only in the pharmaceutical products area. Again, its role in this area is likely analogous to its role in the production of other industrial, commercial commodities whose production is characterized by sharply increasing returns to scale and whose sale is directed to a large number of individual buyers in the form of physicians and to hospitals and hospital-like institutions.

As in the case of advertising, the relative absence of research and innovation as a behavioural characteristic in the major components of the health sector is a cause for concern. Once again, its absence may well be symptomatic of the absence of competitive forces in these areas. This is not to say that there is no research and innovative activity being carried on in these areas. There is. However, what there is, is largely of two kinds. There is a disproportionately large amount of basic research with its potential, long range benefits of world wide applicability. In addition, there is the application of research and innovation carried on in closely associated industries that supply the major and minor equipment for hospitals.

D. Government

To conclude our survey of market conduct, let us deal with government as a separate entity. Government does of course play a role with respect to each of the three major behavioural categories discussed above. There are in addition, however, several characteristics of government's conduct that should be emphasized since each of these is of special interest and warrants specific research enquiry.

Firstly, it has been argued by Buchanan* that a

* J.M. Buchanan, "The Inconsistencies of the National Health Service", Occasional Paper, Institute of Economic Affairs Ltd., November, 1965.

government's entry into the health care sector as the principal funder of health care will follow a particular pattern. In the first instance, it will be concerned with increasing the volume of health care provided in order to meet increases in the quantity of care demanded and any increases in demand that may be a consequence of the participation of government. In a second stage, the population at large, from whose pockets come the tax dollars to support the sector, are subject to a slow but steady process of being educated as to the exact nature of health care. In particular, they begin to see it as a largely private good, whose benefits are almost wholly captured by the recipient of the care, and not a public good the benefits of which accrue to all. In turn, they perceive evidence of misuse and waste by fellow citizens. The third and final stage is one in which a large enough part of the population, having being educated as to the true nature of health care, brings pressure against government and especially politicians to curtail increases in expenditures on health care, if not to actually reduce them.

This Buchanan hypothesis is dealing with an important question, not only with respect to government policy in the health field but also in other fields of economic activity into which it may currently be considering to move. Indeed, there are parts of the health sector, such as the provision of dental care, for which the question is very important.

Clearly, the postulated behaviour of governments and the populations they govern, in combination with the nature of the good in question, should be the object of a continuing research effort. There can be no doubt that different allocative mechanisms, different systems of decision-making, have different specific results.

A second major area of concern about the behaviour of governments is the myopia with which they are so often indicted. For example, in the hospital sector it has been argued that government and especially politicians have been overly impressed with the erection of new hospital structures and correspondingly under enthusiastic about the necessity of operating and maintaining these structures satisfactorily. Again, this ascribed behaviour of governments would seem to warrant a directed and likely continuing research effort. One just doesn't know the facts with any degree of certainty.

The third behavioural characteristic increasingly ascribed to government is related to the nature of its objective function. In particular, it is said of many government departments that they come to act principally as agents for the producer group in question and in so doing harm the cause of the consumer group. Casual evidence, for example, on rapidly rising physician incomes, is consistent with this suggested behaviour. Again, however, this entire question of the objective functions characterizing the provincial and federal ministries of health is one that requires much research

and analysis. Again, we simply do not know.

Market Performance

Our knowledge of the several elements of economic performance in the delivery of health care is rudimentary at best. Thus, in the sections that follow we develop research agendas aimed at filling the more major gaps in our knowledge. Reference to economic principles establishes that not everything is worth the associated cost of knowing. Certainly some things are much more worth knowing than others. In what follows we outline briefly the elements of market performance that seem applicable to the health care sector.

A.1. Technical and Allocative Efficiency: General Summary

As we have noted, of several major world countries Canada has been allocating the largest percentage of its resources to its health care sector. At the same time its health status as indicated by life expectancy, infant mortality, and other measures of mortality is far from the best. Furthermore, we have noted that in one study in which the inputs of the health sector were related to its outputs and in which several important social, economic, and demographic factors were controlled for, Canada was ranked 23rd. out of 25 countries. Though the study was but an initial step in assessing the overall efficiency of the health care delivery system, its results are suggestive of a system that is not as

efficient as it could be. To repeat an old refrain, further research and analysis are clearly called for. Not only money but lives are at stake.

In pursuing this research, it will be necessary to distinguish "technical" efficiency from "economic or allocative" efficiency. The former deals solely with the physical production side of the economic activity. In analyzing technical efficiency, we are therefore interested in assessing whether the production process is technically the most efficient one that might be used given the state of medical science and technology but without regard to the costs of benefits produced.

In contrast, in dealing with economic or allocative efficiency we assume that technical efficiency has already been achieved and enquire if the level of output is appropriate. This question then involves an assessment of the benefits as well as the costs, and in particular involves an assessment of whether the benefits of consuming the last unit of health care produced are greater than or equal to the costs of producing it. In short, have we too much or too little?

A.2. Technical and Allocative Efficiency: Hospitals

The hospital care component of the delivery system has been subject to some study of its technical and allocative efficiency but must still be regarded as an area in which our

ignorance is high. We are perhaps best able to comment on allocative efficiency within this component. At least since the advent of hospital insurance, allocative efficiency is unlikely to have been achieved. Since then, physicians, the principal decision makers, have had almost no incentive whatsoever to inquire, let alone ensure, that the benefits of any procedure requested or performed outweighed the costs of providing it.

In general however, the ultimate assessment of allocative efficiency must be based on a knowledge of both costs and benefits. An analysis of the former is likely to prove more amenable to economic tools and methodology than the latter. We thus concentrate in two following sections on research activity that could be directed to analyzing the technical efficiency of the hospital sector.

A.3. Technical and Allocative Efficiency: Ambulatory Care

As we note in two of the following sections (that dealing with the economics of the group practice of medicine and that dealing with the possibilities for resource substitution in the ambulatory care section) there is a growing body of literature dealing with the efficiency, technical and allocative, of the ambulatory care sector. Much additional work could however be done.

B. Quality

Because of the immense difficulties associated with measuring the output of the health care sector, the quality element of performance takes on special significance. Indicators of health status do of course give some insight into the overall quality of the health system. In addition, there has been some limited study of the quality of individual practitioners of medicine. However, the quality of the goods and services provided in the health sector is but one, difficult to assess, aspect of the generally, difficult to assess output of the sector.

C. Progressivity: Advances in Medical Science and Technology

Only a beginning has been made in recognizing that research and development activities are important to the present and future vitality of the health care delivery system. In one of the following sections we shall discuss this area in greater detail.

As a reflection on available analysis, it is probably fair to conclude tentatively that the hospital sector has often been and importantly so, the site in which the advances in medical science and technology can be introduced. In contrast, the ambulatory care sector, characterized in the main as it is by the independent solo practice of medicine cannot easily be described as being innovative.

D. Equity

Often overlooked, the matter of equity is of course one of the principal justifications for the entry of government into the health care delivery system. It is therefore fair to say that our health care system was not judged to be an equitable one in terms of an equal distribution of health care to those in like categories of need for health. Just what has been the change in the last decade still remains unknown today. Clearly, this is a principal area for research and again one in which the economist may well have an advantage. It is dealt with separately in one of the following sections.

III Outline of the Study

In the next four chapters, we consider the research that might with advantage be carried out on the effectiveness of alternative mechanisms for guiding the principal economic agents in the relevant, broad areas of choice outlined above. We deal successively with incentive systems for physicians, for economic agents in the hospital, for the consumers of health care, and for government in its various forms with special emphasis on the problems inherent in the regionalization of health care resources.

In the sixth and seventh chapters we consider potential areas of research in the non-hospital, ambulatory component of the health sector. We first deal with the organization of the existing resources in this sub-sector. In particular, we address the issue of group practice versus solo practice as the principal mode by which these resources are organized. In our consideration, we take as given the existing institutional arrangements for licencing including the educational requirements, licencing procedures themselves and regulations for various categories of health personnel. We then consider the issue of whether these given volumes of resources, so constrained, could be used more effectively in one mode of organization versus others. In a similar way, we take as given the technology and level of medical science constraining the manner in which non-hospital ambulatory care resources are organized.

The first of these two major aspects of the delivery of non-hospital medical care, assumed given in the discussion of the sixth chapter, is then dealt with separately in the seventh chapter. In it we consider the general possibilities for substituting other types of health care resources for ones currently delivering various subsets of non-hospital health care services. For example, the substitution of the midwife for the general practitioner or obstetrician in the delivery of pre and post natal maternal and child health care; the nurse practitioner for the general practitioner in the delivery of a wide range of ambulatory medical care services; the dental health nurse for the dentist for the provision of dental care to children; the denturist for the dentist for the provision of dental plates, and so on.

In the eighth chapter we broaden our perspective from the ambulatory care sub-sector to that of the entire health care sector and consider the potential for research in the area of the processes by which advances in medical science and technology are introduced into the health care delivery system. We consider not only the potential role of the economist in considering the more specific questions about the introduction of given pieces of new science and technology and their development but also the more general matter of the overall allocation of monies for research and the system by which research is funded.

In the ninth chapter we deal with the potential areas of research on the overall question of the distributive impact

of government funding of the health care services. Though there are likely to be major distributive effects associated with the different areas considered in the first sections of our study, we attempt to deal with these comprehensively in this ninth chapter.

In the tenth, and final chapter we briefly summarize the conclusions reached with respect to the eight major potential areas of research reviewed. In addition, we indicate research opportunities in areas not covered in detail in this study.

It should be emphasized that the selection of issues presented in Chapters II through IX results from the subjective judgment of the author as to which issues of allocative efficiency and distributional equity are likely to be important policy issues in the intermediate run, which are likely tractable given existing and probable data sources and finally which are likely to require the comparative advantage of the economist in their analysis. The selection is thus not designed to be a comprehensive list of all the problems that beset the health sector today. Furthermore, the emphasis given to each of the selected areas was not planned to be equal.

Finally, it was felt strongly that a broadly based menu of research priorities would be much more useful in the Ontario and Canadian context than by a precisely defined ranking because of the wide variety of problems faced by a

wide variety of government bodies and other institutions engaged in some aspect of the funding and/or delivery of health care.

CHAPTER II

The Behavior of Health Care Personnel

I Introduction

As the principal economic actor in the health sector, health care personnel especially the physician and their behavior under different remuneration schemes is a well-justified, principal target for the research activities of economists in the health care field. In a similar way, the behavior of the other major economic actors, namely the hospital administrator, (as integrator of the often conflicting demands of a hospital's medical staff, Board of Trustees and patients), the consumer of health care, and government in its various forms are all likely subjects of research by economists. Though the study of the behavior of each of these actors is highly interdependent, we have chosen to deal with each separately. We begin with the behavior of health care personnel and in particular concentrate on the behavior of the physician. Though direct payments to physicians account for less than twenty-five per cent of the health care dollar, physicians probably play the dominant role in determining how the remaining health care dollar is spent. The physician may thus be considered the principal economic decision maker the health care sector.

In considering the economic behavior of the physician, we first discuss the overall allocation of physician manpower to the health sector. We then proceed to the somewhat more detailed discussion of the allocation of physician manpower within the health sector and will pay particular attention to the role played by different reimbursement schemes. In our discussion we concentrate on the behavior of physicians in the non-hospitals, ambulatory care setting and leave to the following chapter the major discussion of phy-

sician behavior in the hospital setting.

II Overall Supply of Physicians

A. The Blank - Stigler Shortage Model

The study of the overall supply of physicians in a community must almost by necessity involve a consideration of the relative incomes of physicians both in world markets and in alternative occupations in the country in question. From the literature¹ on this subject comes a simple model for examining the relationship between changing incomes and supply. Referred to as the Blank-Stigler Shortage Model, it is illustrated in Diagram II - 1.

Given the implicit assumption that wages for physicians should grow at the same rate as the reference group, for example, all income earners in the community; the new demand function, D_1 , postulated in Diagram II - 1 should in the long run bring about an increase in the supply of physicians such that the new long run equilibrium would be given by the point E_1^{11} with

1. D.M. Blank and G.J. Stigler, The Demand and Supply of Scientific Personnel (New York: National Bureau of Economic Research, 1957).

K.J. Arrow and W.M. Capron, "Dynamic Shortages and Price Rises" The Engineer-Scientist Case, "Quarterly Journal of Economics," Vol. 73 (May, 1959), pp.

W.L. Hansen, "Shortages' and Investment in Health Manpower", The Economics of Health and Medical Care, Edited by S. J. Axelrod (Ann Arbor: The University of Michigan, 1964), pp. 75-91.

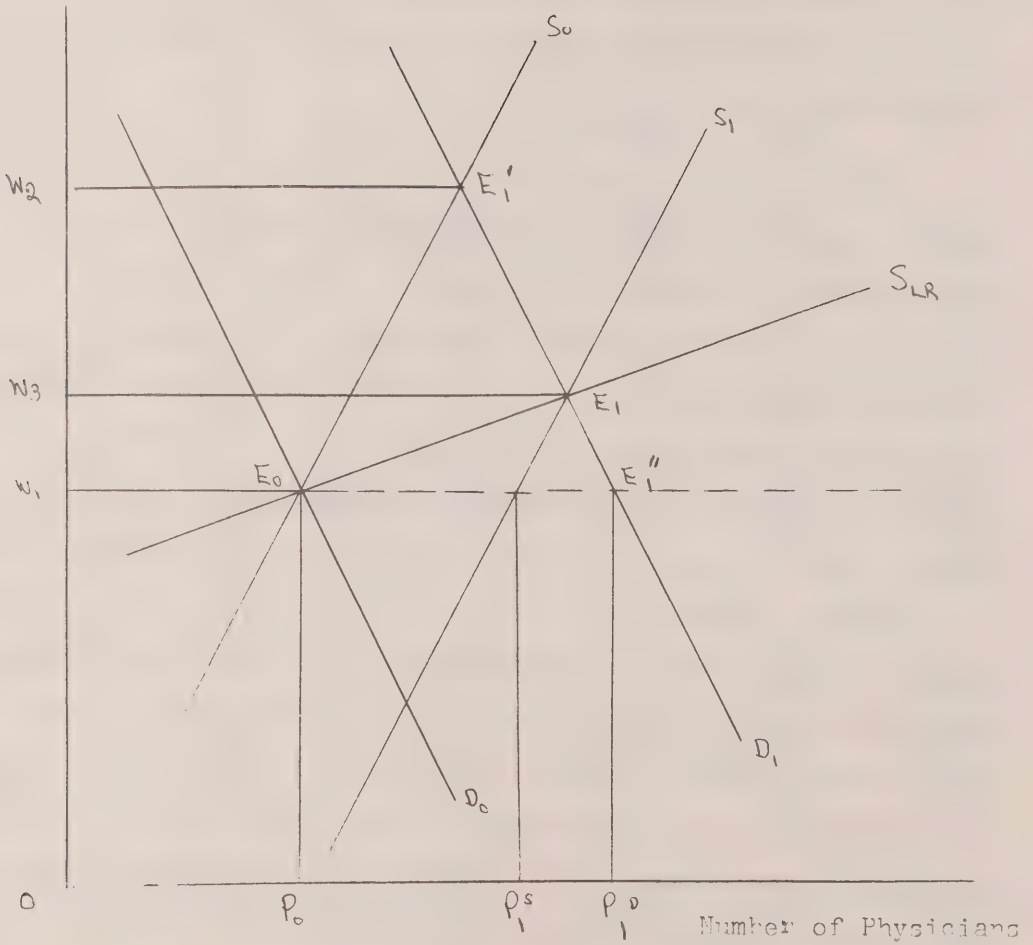
D.W. Black, "The Role of International Flows in the Adjustment of the Canadian Labor Market for Health Professionals", B.A. Honours Thesis (Queen's University, 1967).

D.E. Yett, "The Chronic 'Shortage' of Nurses: A Public Policy Dilemma", Empirical Studies in Health Economics, Edited by H.E. Klarman (Baltimore: The Johns Hopkins Press, 1971) pp. 357-89.

Diagram II - 1

The Blank - Stigler Shortage Model

Relative
Wage



P_1^D physicians supplying their services at the relative wage rate W_1 . If there are constraints to the complete adjustment of supply to the new demand conditions such that the supply function is only able to shift to the position S_1 , then at the relative wage W_1 a shortage of $P_1^D - P_1^S$ is said to exist. Alternatively, if the relative wage line, S_{LR} , that is traced out over time displays a positive slope that cannot be explained by changes in relative conditions of employment, relative costs of education and training, etc., then a shortage can be said to exist.¹

With this shortage model in mind, it is of interest to examine the relative incomes of physicians in Ontario and Canada with respect to other professions and with respect to all taxpayers. In Tables II-1 and II-2 we present data for Ontario and Canada, respectively, on the average taxable income (from all sources) of physicians, "after the expenses of practice have been deducted", and on percentage changes in these incomes for different sub-periods over the years from 1950 to 1972. We present similar information for dentists, lawyers, consulting engineers and architects, all professions, and all taxpayers. Information on the numbers of persons having taxable income and the percentage change in their numbers is also included. Of interest, for example, are the figures displayed in the third last column of Table II-1. The percentage increase in income of physicians over the period from 1950 to 1972 of some 288.5 per cent was accompanied by an increase in the number of physicians of some 143.3 per cent. In contrast, with respect to all taxpayers, the increase in income of 177.4 per cent corresponds to an increase of 196.2 per cent in the number of taxpayers.

1. A crucial assumption is that the initial observations represent equilibrium wage differentials.

Table II-1: Absolute Levels and Percentage Changes in Taxable Income and the Number of Taxpayers for Selected Professional Groups: Ontario, 1950 - 1972.

Occupation	1950	1955	1950-55 % Δ	1960	1955-60 % Δ	1965	1960-65 % Δ	1970	1965-70 % Δ	1970-75 % Δ	1975-80 % Δ
1) Physicians and Surgeons \bar{Y}^* N^{**}	11,032 3,410	14,189 4,130	28.6 21.1	17,563 5,333	23.8 29.1	25,315 5,959	44.1 11.7	39,112 7,630	54.6 28.0	59.2 56.4	122.7 43.1
2) Dentists \bar{Y} N	6,588 1,560	9,114 1,830	38.3 17.3	13,131 1,861	44.1 1.7	17,622 2,238	34.2 20.3	25,242 2,420	43.2 8.1	99.3 19.3	92.2 30.0
3) Lawyers \bar{Y} N	10,906 2,020	15,799 2,330	44.9 15.3	17,256 2,864	9.2 22.9	21,963 3,424	27.3 19.6	32,733 3,652	49.0 6.7	58.2 41.8	49.0 27.5
4) Consulting Engineers & Architects \bar{Y} N	12,682 630	13,183 710	4.0 12.7	15,903 709	20.6 .2	19,102 928	20.1 30.1	22,220 867	16.3 6.4	25.4 12.5	39.7 22.3
5) Total Professionals \bar{Y} N	7,841 11,420	10,103 15,100	28.8 32.2	12,552 18,425	24.2 22.0	16,461 21,367	31.1 16.0	24,287 25,150	47.5 17.7	60.1 61.3	93.5 36.5
6) Total Taxpayers \bar{Y} N	2,971 1,092,520	3,631 1,542,430	22.2 41.2	4,362 1,850,428	20.1 20.0	5,137 2,350,644	17.8 27.0	6,802 3,120,633	32.4 32.3	46.8 69.4	55.9 68.6

* \bar{Y} is Average total income assessed for tax purposes.

** N is Numbers of taxpayers.

Table II-1: Continued

Occupation	1972	1970-72 % Δ	1950-72 % Δ	1960-72 % Δ	1965-72 % Δ
1) Physicians and Surgeons	42,863 8,295	9.6 8.7	288.5 143.3	144.1 55.5	69.3 39.2
2) Dentists	31,124 2,143	23.3 11.4	372.4 37.4	137.0 15.2	76.6 4.2
3) Lawyers	34,886 3,621	6.6 .8	219.9 79.3	102.2 26.4	58.8 5.8
4) Consulting Engineers & Architects	21,496 961	3.3 10.8	69.5 52.5	35.2 35.5	12.5 3.6
5) Total Pro- fessionals	26,502 26,932	9.1 7.1	238.0 135.8	111.1 46.2	61.0 26.0
6) Total Tax- payers	8,242 3,235,810	21.2 3.7	177.4 196.2	89.0 74.9	60.4 37.7

* \bar{Y} is Average total income assessed for tax purposes.

** N is Number of taxpayers.

Table 11-2: Absolute Levels and Percentage Changes in Taxable Income and the Number of Taxpayers for Selected Professional Groups: Canada, 1950 - 1972.

Occupation	1950	1955	1950-55 % Δ	1960	1955-60 % Δ	1965	1960-65 % Δ	1970	1965-70 % Δ	1970-70	1950-60 % Δ	1960-70 % Δ
1) Physicians and Surgeons												
Y*	9,881	12,166	23.1	16,323	34.2	23,229	42.3	24,757	49.6	112.9	65.2	65.2
N**	8,400	10,680	27.1	14,013	31.2	15,410	10.0	19,347	25.5	38.1	66.8	66.8
2) Dentists												
Y	6,202	8,554	37.9	12,238	43.1	15,693	28.2	22,794	45.2	86.3	97.3	97.3
N	3,570	4,020	12.6	4,381	9.0	5,357	22.3	5,665	5.7	29.3	22.7	22.7
3) Lawyers												
Y	9,641	12,243	27.0	14,597	19.2	19,191	31.5	26,738	39.3	83.2	51.4	51.4
N	4,550	5,570	22.2	7,195	29.2	8,374	16.4	9,304	11.1	29.3	58.1	58.1
4) Consulting Engineers & Architects												
Y	10,955	14,007	27.9	15,670	11.9	19,279	23.0	22,385	16.1	42.9	43.0	43.0
N	1,500	1,870	24.7	2,019	8.0	2,767	37.0	2,483	10.3	22.3	34.6	34.6
5) Total Professionals												
Y	7,204	9,316	29.3	12,135	30.3	15,665	29.1	21,911	39.9	80.6	68.4	68.4
N	24,140	35,650	47.7	44,418	24.6	52,134	17.4	62,577	20.0	40.8	84.0	84.0
6) Total Taxpayers												
Y	2,962	3,535	19.3	4,232	19.7	4,947	16.9	6,447	30.3	52.3	42.8	42.8
N	2,374,240	3,558,650	49.9	4,389,766	23.4	5,728,942	30.5	7,641,731	33.4	74.1	84.9	84.9

* Y is Average total income assessed for tax purposes.

** N is Number of taxpayers.

Table II-2: Continued

Occupation	1972	1970-72 % Δ	1950-72 % Δ	1960-72 % Δ	1965-72 % Δ
1) Physicians and Surgeons	41,194 22,065	18.5 14.0	316.9 162.7	152.4 57.5	77.3 43.2
2) Dentists	28,362 5,097	24.4 10.0	357.3 42.8	131.8 16.3	80.7 4.9
3) Lawyers	30,603 9,300	14.5 -----	217.4 104.4	109.7 29.3	59.5 11.1
4) Consulting Engineers & Architects	25,477 2,412	13.8 2.9	132.6 60.8	62.6 19.5	32.1 12.8
5) Total Pro- fessionals	26,396 2,412	20.5 5.7	266.4 174.0	117.5 48.9	68.5 26.9
6) Total Tax- payers	7,784 8,071,015	20.7 5.7	162.8 240.4	83.9 84.1	57.3 41.1

* \bar{Y} is Average total income assessed for tax purposes.

** N is Number of taxpayers.

Source: Canada, Department of National Revenue, Taxation Statistics, 1952 through 1974.

The information on relative incomes is then presented directly in Tables II - 3 and II - 4 for Ontario and Canada, respectively. We note clearly the significant jumps in the relative incomes of physicians from 1960 to 1965 and then again from 1965 to 1970.¹ Without question, a prime research project would be one whose objective was the explanation of the differences in the relative wages paid to the different professions in terms of employment conditions, education and training costs and so on. Alternatively, this research might be thought of as the testing of different hypothesis about the impact on the relative, income-numbers relationship of such changes in circumstances as: the introduction of the national-provincial hospitalization plans in the 1959-1961 period, the introduction of the medical plans in the 1965 to 1969 period, the sharply changing balance between general practitioners and specialists, inter-provincial and inter-national movements of physician manpower, the number of places in Canadian and Ontario Medical school, and so on. The impact of each of these factors would be an interesting and worthwhile study in its own right. More fundamentally the existence and nature of the relationship between relative incomes and numbers of physicians is prime subject matter for economic research.

Closely related is the issue of the relationship between overall supply and the distribution of medical manpower by geographic location, economic and social class, demographic factors and so on. Similarly important is the determination of a satisfactory overall level of income.

1. A word of caution should be given in the use of these figures. Because they derive from taxation statistics they include income from all sources, not just that of the recipients principal occupation. Secondly, they exclude those income earners who do not generate enough income to pay taxes. It would be of some interest to compare the picture set out above with that describable with data on incomes and employment taken from the population census.

Table II-3: Average Income* of Selected Professional Groups Relative To That of All Taxpayers: Ontario, 1950 - 1972.

Occupations	1950	1955	1960	1965	1970	1972
1) Physicians and Surgeons	3.7	3.9	4.0	4.9	5.8	5.2
2) Dentists	2.2	2.5	3.0	3.4	3.7	3.8
3) Lawyers	3.7	4.4	4.0	4.3	4.8	4.2
4) Consulting Engineers and Architects	4.3	3.6	3.6	3.7	3.3	2.6
5) Total Professionals	2.6	2.8	2.9	3.2	3.6	3.2

* Average total income assessed for tax purposes.

Source: Table II-1 and in turn Canada, Department of National Revenue, Taxation Statistics, 1952 through 1974.

Table II-4: Average Income* of Selected Professional Groups Relative To That of All Taxpayers: Canada, 1950 - 1972.

Occupations	1950	1955	1960	1965	1970	1972
1) Physicians and Surgeons	3.3	3.4	3.9	4.7	5.4	5.3
2) Dentists	2.1	2.4	2.9	3.2	3.5	3.6
3) Lawyers	3.3	3.5	3.5	3.9	4.2	3.9
4) Consulting Engineers and Architects	3.7	4.0	3.7	3.9	3.5	3.3
5) Total Professionals	2.4	2.6	2.9	3.2	3.4	3.4

* Average total income assessed for tax purposes.

Source: Table II-2 and in turn Canada, Department of National Revenue, Taxation Statistics, 1952 through 1974.

This last issue

is of course intimately connected to the question of the optimal numbers of physicians. If optimal numbers were established, one could adjust incomes to bring forth the appropriate supply. Without the setting of an optimum number to guide the establishment of levels of income, we are left in the more nebulous state of being concerned that physicians be paid a wage rate in keeping with their relative conditions of employment, relative costs of education and training, etc. For example, if a relative wage of four times that of the average taxpayer sufficed to call forth what is judged to be a reasonable supply of medical manpower over the period from 1950 to 1960, is there sufficient justification for the fact that a relative wage in excess of 5 is required in the 1970's?¹ Finally, the impact of constrained physician supply on the ability of physicians to obtain disproportionately high increases in average income is a worthwhile area of research. For example, is it the case that the situation of constrained supply in the face of growing demands by the community provides the political environment in which physicians are more easily able to achieve "disproportionate" increases in income through negotiations with government on the level of the increase in their fee schedule?

1. A literature is fast growing on the subject of the real returns to a medical education and the corollary what is a reasonable income. For example, see C.M. Lindsay, "Real Returns to Medical Education," Journal of Human Resources, Vol. 8 (Summer, 1973), pp. 331-48. See also F.A. Sloan, "Life-time Earnings and Physicians' Choice of Specialty", Industrial and Labour Relations Review, Vol. 24 (October, 1970) pp. 47-56.

B. Foreign Trained Physicians and the Medical Market

The recently announced decision by the Federal government to control the flow of foreign trained physicians provides an interesting case study for further exploration of the market for physicians services. In particular, we wish to enquire whether there is in the present system a mechanism for controlling the overall supply of physicians services and the price at which they are supplied. We attempt to demonstrate that no such mechanism currently exists and therefore that this market is at present a non-equilibrating market.

As an empirical base, we note the data shown in Table II - 5 on the sources of Ontario physician manpower. The role played by foreign trained

Table II - 5 - Ontario Physician Manpower Sources, 1968 - 1972.

(By year of initial registration)

Year	Ontario Medical Schools		Other Canadian		Foreign		Regular Register Total	
	No.	%	No.	%	No.	%	No.	%
1968	278	45.4	116	19.0	218	35.6	612	100
1969	284	38.8	128	17.5	320	43.7	732	100
1970	340	36.9	200	21.7	382	41.4	922	100
1971	342	35.9	158	16.6	452	47.5	952	100
1972	359	35.1	164	16.0	501	48.9	1,024	100

Source: College of Physicians and Surgeons of Ontario, Registration Statistics, December 31, 1972, as printed in the Report of the Health Planning Task Force, (Province of Ontario, 1974) Table 7,1, p. 65.

physicians is seen to be a major role, as they accounted for one half the number of new registrants in 1972. It is also an increasingly significant role. Without having detailed information on relative prices for physician's services in the different countries from which these graduate emigrate and Canada, it is difficult to be precise about the reasons for this flow. However, the existence of a world market for physicians services has been amply demonstrated previously.¹ It thus seems fair to entertain as a principal hypothesis the likelihood that relative fees per unit of service are significantly higher in Canada and especially in Ontario and that as a result the flow has been in the direction and of the magnitude that we have just indicated. Clearly other factors such as job conditions, complementary facilities and manpower and so on must also be considered.

An interpretation of the impact of this flow on the market for physicians services in Canada may now be briefly sketched as is done in Diagram II - 2. The familiar price-quantity axes refer to units of service provided by the physician where the principal criterion for determining the unit is the time involved in providing the service; it is then assumed that the "complexity" of the service and any disproportionate requirements for the help of auxiliary personnel and the input of capital plant and equipment are

1. D.W. Black, "The Role of International Flows in the Adjustment of the Canadian Labour Markets for Health Professionals", B.A. Honours Thesis (Queen's University, 1967).

Consistent with this market model within the confines of Canada are the data examined by R.G. Evans. Indeed he states that "relatively high physician incomes in a province tend to draw in physicians but only weakly and with a lag. More rapid increases in physician stock seem to moderate the rise in physician incomes, implying either that prices rise less rapidly or that services rise less rapidly (or both). Average physicians incomes do not seem to be related either to a price adjusted trend or to incomes in the community around them...." R.G. Evans, Price Formation in the Market for Physicians Services A Study Prepared for the Prices and Incomes Commission (Ottawa: Information Canada, 1972), p. 73.

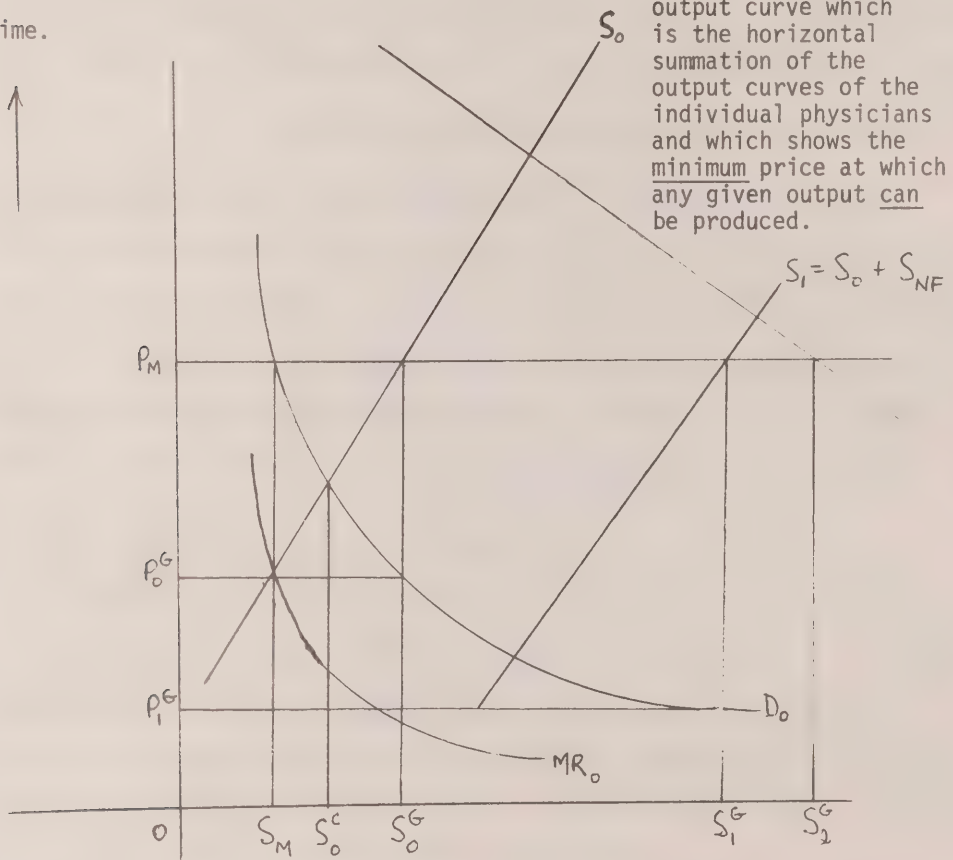
Diagram II - 2

The Market for Physicians: An Unstable Equilibrium

Price per

unit of service

time.



#of units of service /t
(adjusted for quality and,
or complexity of service in
question; it is assumed that
OMA is roughly successful in
adjusting fees payable for
different services to reflect
principally the time and effort
involved and capital facili-
ties and equipment.

also taken into account.

As a first step, the determination of price and quantity in the pre-government insurance plan era is assumed to follow monopoly pricing policy with price therefore set at P_M and quantity supplied at S_M . The short-run output function, S_O , is probably fairly price inelastic whereas the demand curve, D_O , probably is fairly price inelastic over the initial range of output but thereafter becomes increasingly price elastic.

The second step portrays the breaking of the link between quantity supplied and quantity demanded at market price with the introduction of government funded insurance. Price to the physician, the supplier, is given by P_M (strictly speaking 90 per cent of P_M) regardless of the level of output; P_M is thus the effective average revenue and marginal revenue for the physician. He is thus led to provide S_O^G services.

For his part the consumer now faces no direct price, (or at the limit 10 per cent of the price P_M),¹ and thus his effective demand is given by a position on his demand curve in its most elastic portion. In fact, his demand is limited only by the indirect costs imposed on him of seeking health care where such costs include an evaluation of his time, transportation costs and so on.

What is of further interest however is that at the output level, S_O^G , the real cost of the last unit in production, namely P_M , is substantially in excess of the value of that last unit to the consumer, namely P_O^G .

1. Physicians who have opted out of the common reimbursement procedure and who bill their patients who in turn seek reimbursement from government are billing their patients directly.

Now if the price P_M is indeed such as to permit the generation of annual incomes significantly in excess of those possible in other countries, a flow of foreign trained physicians will be induced. Such a flow will shift the output curve to some new position S_1 . The new level of output S_1^G , will then exhibit an even greater discrepancy between the real costs of producing the last unit of care and the value of that care to the consumer.

There are two essential features of this market model that require emphasis. First, the flow of additional physicians does not require that either new and or existing physicians are able to significantly affect the demand for their services.¹ As long as we assume that the wants of individuals with respect to physicians services are proximately insatiable, for example, as shown by the elastic portion of the demand curve in Diagram II - 2, then the Canadian Medical Market could support a very large (infinite?) number of physicians, given the government funding mechanism.

Secondly, the market in the insurance era is a non-equilibrating one compared with that which existed in the preceeding period. In contrast to the current situation, in the preceeding one the flow of new physicians to Canada in response to the relative high price would likely have ultimately led to a decrease in the monopoly price and thereby have brought the Canadian market slowly into equilibrium with the world market.

As an aside it is interesting to speculate on the impact of these large numbers of foreign trained physicians on the level of the fee-for-

1. Relaxing this strong assumption would of course lead to even greater discrepancies between the real value and real costs of services rendered.

service schedule. At present in Ontario the absolute level of fees and, in principle, the relative level of fees within the schedule are subject to negotiation between representatives of the physicians and those of government through the medium of the Joint Committee on Physicians' Compensation. Though the workings of this committee cannot be precisely set forth, it is likely the case that the existence of strong public sentiment as to the perceived "shortage" of physicians might well work in favor of relatively larger increases in fees. In other words, had the flow of foreign trained physicians been subject to greater restrictions over the last five to ten years, the overall level of physician fees might be substantially higher than it is today.

Probably the principal qualification that should be advanced to the interpretation advanced above involves a consideration of the destination of the foreign trained physicians. To the extent that they locate in low income rural areas that are currently poorly serviced by physicians, the value of physician care might more closely approximate the costs of its provision if indeed not even exceed it; that is, the demand for care by persons living in such areas might well differ markedly from that describing the general population, as discussed in Chapter X, The Distributional Impact of Government Funded Health Care.

Data describing the location decisions of foreign trained physicians are not readily available. Anecdotal evidence suggests that physicians practicing in poorly serviced areas are frequently foreign trained. At the same time, similarly gathered evidence suggests that many foreign trained physicians are led in their attempts to gain a licence to practice in Canada to

become residents in hospitals and that further a disproportionate number of them become involved in the speciality training programmes. Once again, hard evidence is lacking.

An alternative interpretation to the one we have just advanced is that the demand of Canadians for physicians' service is very strong indeed and that it occupies a position such as D_2 in Diagram II - 2 relative to the current supply S_1 . In this case, a different set of conclusions would be forthcoming about the welfare implications of restricting the supply of foreign trained physicians.

In the foregoing we have noted the several bodies of data that do not currently exist in a readily usable form and some of the hypotheses that might be examined if such data were available. There is one principal issue that we have not dealt with but yet one that must be placed on the record, namely, the overall judgement as to how much physician care is "enough". One can rationalize a government's decision to restrict supply as an attempt to approximate the level of output S_1^C given that D_0 and S_1 are the relevant demand and supply curves. On the other hand, speaking of a generalized demand curve such as D_0 is probably disastrous as far as portraying the actual demands of individuals for care. For the demand for care is seen in the limited studies that have been done to date to be highly variable. If this is indeed the case, then we ultimately want to establish a system in which the demands of individuals can be brought to bear in the market especially if they are in excess of what are deemed to be either average or necessary demands. In other words, should not the ultimate objective be to establish a system in which the government can determine how much care it wishes to buy for

individuals and also that individuals be free to vary their consumption in accordance with their own priorities. If we accept this as a goal, of a guaranteed minimum then the determination of the welfare implications of the flow of foreign trained physicians, as laid out above with the aid of Diagram II - 2, is clearly inadequate.

III Allocation of Physician Manpower Within the Health Care Sector

A. Introduction

Though in principle one cannot separate decisions on the allocation of physicians' activity within the health sector from those on the overall allocation of manpower to the sector itself, we will do so in order to concentrate our attention in this section on the allocational effects within the sector of alternative reimbursement schemes. Our principal justification for doing so is that most of the schemes under consideration permit one to establish the overall size of the payments to physicians and thereby to roughly determine the average income of physicians for a given number of them or alternately to vary average income in order to affect the number of physicians.

In Figure II-1, we have set forth the two broad alternative methods by which government can finance health care, "tiers garant" and "tiers payant", and in each of these classes the principal variants that have been used or proposed for use in different health care delivery systems in the world. With respect to the alternatives for physician reimbursement under "tiers payant", direct control of average physician incomes is in principle easily done if the remuneration scheme is based on salary or on capitation. Similarly, with fee-for-service and care of episode payment schemes, a fixed pool of monies can be established each year and rates or fees pro-rated to exhaust

the pool each year.

Figure II - 1. - Alternative Methods of Financing Health Care

A. "Tiers Garant"

Government Reimbursement of Patient

Full	Fixed*	Sliding**	Fixed*	Sliding**
	%	%	Deduc- tible	Deduc- tible

B. "Tiers Payant"

Direct Government Payments to Providers

Fee-for- Service	Case Payment	Capitation	Salary	Global Ambulatory	Global Hospital	Global Ambulat- ory and Hospital
---------------------	-----------------	------------	--------	----------------------	--------------------	---

* Patient could of course be required with either fixed percentage or fixed deductible to pay out-of-pocket the first dollar of the bill for each contact with the physician or the first ten dollars of the bill for each day in hospital.

** Sliding percentage or deductible is geared to taxable income.

Source: See text.

Slip year funding or a thirteenth payment based on the experience during the twelve months of the year are obvious ways of arranging for the exhaustion of the fixed pool. With the global budgets to cover a group of physicians either in the hospital or ambulatory care setting separately (or the two taken together) the direct control of the incomes of physicians would be somewhat more difficult in principle to effect. In practice

however, with the wages of other personnel fairly well known as well as the costs of supplies, etc. it would likely prove fairly easy to directly influence the average level of physician income.

With respect to the "tiers garant" class, direct control of physician incomes would only be possible in those variants that permit both the fixing of the total pool of monies to be spent by government each year and the setting of a legal requirement that physicians charge the patient a fee that varies within some fixed percentage of the amount with which the patient will be reimbursed. In general, it would seem to be more difficult to control physician incomes under the "tiers garant" class of reimbursement schemes.

We should emphasize however that it is the "tiers payant" class of remuneration schemes that encompass the set of alternatives being considered currently.¹ With this class, it is possible to directly control average incomes, given the number of physicians, or alternately to control numbers through varying income, for each variant except those involving a global budget and in these it would likely be fairly easy to do so. We are thus able to concentrate in this section on the impact of alternative remuneration schemes on the allocation of physician activity within the health care sector. We first review the criteria by which these schemes might be judged; we then present a brief summary of the literature in the area; and finally, we set out a brief statement of the pros and cons of four alternative schemes; salary,

1. For example, The Report of the Health Planning Task Force argues that the "tiers garant" schemes are not acceptable, Report of Health Planning Task Force (Toronto: Province of Ontario, 1973), Chapter 6, "Financing of Health Care", pp. 57 - 58.

capitation, case of episode payment, and fee-for-service.

B. Criteria for Evaluating the Efficacy of the Remuneration Scheme

The list of criteria for evaluating the relative goodness of a particular remuneration scheme is a long one. In addition to the criterion discussed in the previous section, namely the impact of the scheme itself on the average level of income and thereby on the numbers of physicians and those striving to become physicians, the principal criteria would appear to be the following: the degree of abuse likely or alternately the costs of enforcing the adherence of physicians to the scheme, the extent of incentives to provide the appropriate quality of care, the extent of incentives to provide the appropriate quantity of care, the flexibility of the system in facilitating adjustments to changes in the demand and supply of health care, the likely size of administrative costs involved in attempting to match remuneration to the real social value of the services rendered, the ability of government to provide incentives for physicians to emphasize particular kinds of health care, such as preventive care, that they might otherwise be inclined to neglect, the impact of the scheme on the socio-economic, demographic and/or geographic distribution of the care received, the extent of incentives to adopt and use the appropriate level of "advances" in the art and science of medical care, and the likely costs involved in changing from the existing method of remuneration to a different one.

Many of the above criteria have been raised in the more general discussion of the reasons underlying the formation of "firms" and their use of

managerial "fiat" for directing activities within their jurisdiction rather than relying on individuals, as solo producers, and the use of the price mechanism to co-ordinate their individual economic activities. Whether in the guise of enforcement costs, transactions costs, or simply information costs, this literature, including the work of F.H. Knight, R.H. Coase, A.A. Alchian and H. Demsety, and J. McManus,¹ focuses on the problems involved in ensuring that the activities and/or, goods and services being purchased are the same as those being supplied through use of the price system as compared with the organization of economic activities by the firm using fiat as the principal means of directing economic agents. In the context of the choice of remuneration schemes for physicians, the frameworks discussed in the literature just mentioned seem to emphasize the need to consider on the one hand the problem of over-production when prices are used as in the case of fee-for-service and case or episode payment and of under-production when separate activities are not priced as in the case of capitation or salary.

1. F.H. Knight, *Risk, Uncertainty and Profit* (New York:

R.H. Coase, "The Nature of the Firm", *Economica*, Vol. 4 (November, 1937) pp. 386 - 405; reprinted in G.J. Stigler and K. Boulding, (Eds.) *Readings in Price Theory* (Homewood: R.D. Irwin, 1952), pp. 331 - 51.

A.A. Alchian and H. Demsety, "Production, Information Costs, and Economic Organization", *American Economic Review*, Vol. 62 (December 1972), pp. 777 - 795.

J. McManus, "The Cost of Alternative Economic Organizations", (Unpublished, 1973).

C. Literature Survey

Hogarth¹

J. Hogarth's work on comparative systems of paying physicians in publicly organized or sponsored health care programs is one of the most detailed descriptive studies available on the subject. Hogarth concentrates his efforts on the United Kingdom, Germany, Austria, France, Switzerland, Holland and the Scandinavian countries. Also included are less detailed descriptions of the payment schemes in New Zealand, Australia and the Soviet Union. The detailed lists of the advantages and disadvantages of the schemes found in each of the different countries does not, however, lead to generalizations that would be clearly valuable in the determination of the optimal remuneration scheme in any particular country. What we know as a result of Hogarth's work is that the physician, the patient, and the organization which finances the health care program each have a set of objectives (some of which are in conflict) that each hopes the remuneration scheme will accomplish. The extent to which the objectives of any one group are achieved in a particular country is a function of many factors.

Glasser²

Following along the lines of Hogarth's work is that of W.A. Glasser in his study of the remuneration schemes in sixteen countries in Europe, the Middle East and the Soviet bloc. Glasser is more interested that Hogarth in

-
1. J. Hogarth, The Payment of the Physician (New York: Pergamon Press 1963).
 2. W.A. Glasser, Paying the Doctor: Systems of Remuneration and their Effects (Baltimore: The Johns Hopkins Press, 1970)

the political circumstances surrounding the establishment and development of the different schemes. Once again however it is difficult to deduce clear-cut policy options from the detailed analysis and evaluation of the many different remuneration systems studied.

Alexander¹

In 1963 the method of payment for physicians in the Medical Care Program for the Indigent operated by the City of Baltimore was changed from capitation to fee-for-service. C.A. Alexander has analyzed the impact that this change had on the utilization of physician services, clinics and hospitalization. Since no other significant changes in the scheme occurred, it was possible for Alexander to set out reasonably comparable statistics. Among his findings are the following: the average number of visits to private physicians operating in the plan increased from a mean of 2.65 visits in 1961 and 1962 to 3.1 visits in 1963 and 1964; visits to medical care clinics seemed to remain fairly constant; the average number of prescriptions per patient increases from 5.3 to 6.35 from 1961 - 62 to 1963 - 64; and the average number of in-patient days of hospital care appeared to have remained relatively stable. Administration costs of the program excluding those for in-patient care fell from 7 to 5.8 percent of the total costs of the program over the period from 1961 to 1964, largely because of the discontinuation of the sampling procedures used with the capitation scheme to determine the utilization of physician services under that scheme.

1. C.A. Alexander, "The Effects of Change in Methods of Paying Physicians: The Baltimore Experience", American Journal of Public Health, Vol. 57 (June, 1965), pp. 1278 - 89.

Because of the lack of data describing changes in utilization and costs in the Baltimore community at large over the period 1961 to 1964, it is difficult to derive precise conclusions about the impact of the method of payment on costs and utilization. Furthermore, there is the absence of a standard of what is the "appropriate level" of care with the result that proponents of capitation can argue that fee-for-service led to over-production while at the same time proponents of fee-for-service can argue that capitation led to the neglect and under-production of services for the group of indigents in question. One factor that Alexander's study helps to emphasize is that under capitation, as indeed under salary, some mechanism will likely be required for monitoring the number and nature of the activities carried out by physicians. Under fee-for-service, such information is a by product of the billing system.

Monsma¹

Using data from a survey of approximately 400 families in each of the Health Insurance Plan of Greater New York (HIP) and Group Health Insurance, Inc. (GHI) as reported by O. Anderson and P. Sheatsley, G.N. Monsma explored three hypotheses related to the interdependence of demand and supply for health care. They were as follows:

1. that the higher the marginal revenue to the physician for performing a given service, the greater the demand for it will be;

2. this increased demand will be greater for surgical procedures (and other procedures that require hospitalization) than for home and office visits...since illnesses requiring this type of treatment are generally more serious and a patient is more reluctant to disregard his doctor's advice;

1. G.N. Monsma Jr., "Marginal Revenue and the Demand for Physicians' Services", Empirical Studies in Health Economics edited by H.E. Klarman (Baltimore: Johns Hopkins Press, 1970), pp. 146 - 58.

3. the increased demand for surgery will be concentrated among those procedures that involve the removal of organs, the absence of which will not greatly impair the functioning of the individual and for which the need for the procedure is subject to some doubt (for example, tonsillectomies, appendectomies, and hysterectomies)."¹

It should be noted that under GHI the marginal revenue is clearly positive since physicians are remunerated on a fee-for-service basis whereas under HIP with capitation marginal revenue is zero, at least in the short run. Monsma demonstrates that with respect to non-surgical, non-obstetrical procedures, there is little difference between the two groups. However, with respect to hospital based surgical procedures, the GHI group had 7.2 procedures per hundred per year performed on them compared to only 4.4 for the HIP group.

These results of Monsma clearly seem to indicate that physicians decisions on which treatment packages to provide patients with given diagnostic problems are a function of the fee that can be earned by providing each treatment package and that this relationship is in turn a function of the nature of the diagnostic problem. Given the absence of an agreed upon standard of what constitutes the appropriate number of surgical procedures for an average population, it is, however, difficult to conclude either that the GHI group was over-provided with such care or that the HIP group was under provided.

Marmor and Thomas²

1. Ibid., p. 149.

2. T. Marmor and D. Thomas, "The Politics of Paying Physicians: The Determinants of Government Payment Methods in England, Sweden, and the United States", International Journal of Health Services, Vol. 1, Number 1 (1971), pp. 71 - 78.

Marmor and Thomas are concerned with testing the hypothesis that "whatever the political and medical structure of a western industrial country, physician preferences determine the governmental methods of payment".¹ The underlying model on which this hypothesis is based is the following:

Premise 1. Doctors in western industrial countries prefer payment methods in public programs with which they were familiar before the onset of the public program in question.

Premise 2. Doctors are presumed to be willing to strike over government efforts to change these familiar payment methods....

Premise 3. Western industrial states will never risk a medical strike because of the high political costs associated with the interruption of personal health services irrespective of government views on the merits of physician demands concerning payment methods.

Premise 4. Such governments...prefer gaining medical concessions on the amount of expenditures in exchange for concessions on methods of payment.

Premise 5. The failure to satisfy widely understood medical preferences on payment methods is presumed in western industrial countries to be the sufficient condition for a physicians' strike.

Premise 6. In general, government medical officials prefer salary method of payment."²

The evidence gathered to test their hypothesis included a detailed investigation of the experience of England, Sweden and the United States since World War II as well as less intensive study of decisions on payment methods in several other industrialized countries including the Netherlands, West Germany, France, Switzerland, Spain, Italy, Canada, Greece, Poland, the Soviet Union and Israel.

The general conclusion of the authors is that "the doctors get their way on the methods of their pay!" This, they argue, is due in the main to the

1. Ibid., p. 72

2. Ibid., pp. 73-74.

significant political-economic resources of physicians relative to the value placed by governments on the benefits likely to accrue should payment methods be changed and in turn their decisions as to whether to allocate the necessary resources to effect such a change. Though the authors do not provide us with information on the costs of changing methods of payments in a way that would allow us to balance against them perceived benefits from such a change, their evidence does emphasize that such costs would appear to have been especially significant as revealed by the behavior of governments in not pursuing their objectives to change payment methods.

D.- Some Tentative Propositions of the Impact of Four Reimbursement Schemes

With respect to the four schemes, salary, capitation, case or episode payment, and salary, the following are some tentative propositions that can be deduced from the literature in this area. They are set forth on the assumption ceteris paribus, especially mode of practice.

Abuse

All reimbursement systems are subject to abuse by patients and physicians if they are without appropriate mechanisms for ensuring financial and programme accountability. This is so even in a free market system with the use of fee-for-service since physicians generally possess some degree of monopoly power and because the quality of the service rendered is so difficult to judge for the consumer. Even if no such power existed, the costs of gathering information about the quality and/or quantity of the good or service produced would mean that some level of abuse was "optimal"; that

is, the costs of eliminating all abuse would exceed the benefits of doing so.

Income Limits

The average income of a physician when directly funded by government can be controlled under any reimbursement scheme. Under the fee-for-service and case payment methods, a fixed pool of monies can be established and the rates of reimbursement pro-rated to exhaust the pool. The capitation and salary methods provide more obvious ways of fixing income levels.

Incentives for Quality

Incentives to provide a high quality of care are only weakly affected by any reimbursement system. Factors other than the reimbursement scheme itself seem to bear more significantly on this aspect. To the limited extent the reimbursement scheme does provide an incentive for high quality care, it is likely to be highest with fee-for-service, followed successively by case payment and capitation. This proposition is based on the premise that the flow of patients to a physician may be in part a function of his reputation for providing high quality care (to the extent such a reputation is known and knowable). One can of course conceive of changes being brought about in the knowledge available to consumers about the competence of individual practitioners that would change the current situation. In addition, one can conceive of extensive monitoring schemes being used in salary or capitation schemes that might alter the proposition advanced above. Other things being equal, however the proposition is probably correct in the present situation.

Incentives for Quantity

Incentives of varying import to provide a high volume of care are inherent in three systems of reimbursement but non-existent in a fourth, namely the salary system. Once again, the mechanism for gathering information on the basis of which to determine merit differentials in salary or in the unit of value of services (under a capitation scheme) might well narrow significantly the differences expected; other things being equal.

Short Run Supply and Demand Adjustments

Short-run adjustments in the supply of services to meet changes in demand are probably most easily accomplished with the fee-for-service system, followed closely by episode payment, and at a greater remove by capitation. These adjustments are likely to be slowest under the salary method; such demand changes include both those that affect the overall demand and those that affect the composition of demand.

Administrative Costs and Financial and Program Accountability

In order to ensure financial and programme accountability administrative costs borne by the physician and government ministry are likely to be high in all cases. They are nevertheless likely to be lowest with capitation followed by case payment and fee-for-service, respectively. Administrative costs for the salary method are very much a function of the variation in salaries thought equitable because of differential levels of productivity. These variations in turn might be a function of the volume of services provided. As a rough estimate, the salary system most likely to be chosen,

including the method of determining and upgrading salaries, would involve administrative costs similar to those of a fee-for-service system.

Reimbursement Schemes and use of Complementary and Substitute Resources

The impact of different reimbursement schemes on the use of complementary and substitute health resources is unclear. On the one hand, resources when considered as complements to the physician's services might well be used more widely and perhaps "overused" with the fee-for-service method and correspondingly "underused" under capitation and salary methods. On the other hand, when considered as substitutes for physician services, resources might well be underused under fee-for-service and correspondingly overused with capitation or salary. Factors other than the basic reimbursement scheme itself might thus be more important in ensuring a rational use of alternative health resources, including hospital beds, para-medical personnel, pharmaceutical products, etc. As has been emphasized earlier, the absence of commonly agreed upon standards of health care including indicators of health status precludes estimation of the extent of over or under production under the various methods.

Relative Emphasis on Preventive Care

Because of the inherent nature of the fee-for-service method, it should be possible to more directly influence the provision of preventive health services with it than under either capitation or salary; that is, by increasing the fees payable for preventive services, the activities performed by a physician can be directly influenced. In contrast, under salary or capi-

tation it would seem somewhat more difficult to emphasize preventive care. Recourse must be had under these schemes to educational programmes or to monitoring systems which measure in detail the volume of preventive care provided.

Impact on the Socio-Economic, Demographic and Geographical Distribution of Care

With regard to the geographic distribution of care, subsidies in addition to the general reimbursement package seem necessary to provide health resources in centers other than the major urban settings. It is not clear whether the nature of the general reimbursement scheme bears significantly on the ability of the province to provide these subsidies.

With regard to the socio-economic and demographic distribution of care, the fee-for-service method would seem to provide the greatest incentive to physicians to provide care to the disadvantaged where disadvantageousness is defined in terms of social background, age, etc. It does so by providing a direct financial incentive to the physician to provide care to anyone regardless of his age or socioeconomic background.

Incentives for the Use of New Technology

There appear to be two prime types of technical progress those that relate to advances in medical science and those that relate to improved methods of organizing the delivery of care as in the use of paramedical personnel, etc. With regard to the first component, fee-for-service would seem most likely to provide the physician with the strongest incentive to introduce the new technology. Case payment, capitation, and salary would follow

in that order. With regard to the second component, the same ordering is again plausible although in this case the preferences of physicians for a particular salary system is often accompanied by strong preferences for various types of practice and thus observations from real life tend to obscure the impact of the reimbursement scheme itself on incentives for innovation.

The above propositions are highly speculative and follow from a subjective evaluation of the research in the area that has been carried on to date. They are advanced in an effort to provide a starting point for future research; to adumbrate the issues. Can they be supported or rejected more firmly as a result of further research?

In setting out these propositions there were two major pitfalls that we attempted to avoid. The first of these relates to the confusion that often exists between the relative merits of a particular remuneration scheme "if operated in a reasonable manner" and the merits of the scheme if operated otherwise. For example, in considering the merits of a fee-for-service scheme for Ontario one would want to visualize its operation with a reasonable set of checks on the billing practices of physicians. For the first years of the operation of the fee-for-service scheme within the medical care programme, there was almost a complete absence of such checks. Had physicians during this time been paid under capitation or salary it is difficult to see why their incomes would not have risen at the same pace, (unless salaries might be said to be more visible). In other words for the past several years, Ontario and Canadian society appears to have been prepared to accept the outcome that relative physicians incomes should rise dramatically.

The second pitfall to be avoided relates to a concern for the environment in which the physician operates. The costs of enforcing different remuneration schemes are likely to be significantly affected according to whether a physician provides his services in relatively isolated, solo practice or whether he works as part of a health care team in either an ambulatory or hospital care setting. In the solo practice setting, the costs of enforcing a salary system are likely to be relatively high compared to what they might be in the latter setting. Just the reverse is likely true for the fee-for-service setting.

IV Research Agenda

There would seem to be at least four major objectives that the research of economists might pursue in this rather broad area of the behavior of physicians in response to alternative remuneration schemes. First, with respect to the relationship between the average level of income and the supply of medical manpower, there are a number of highly interdependent issues that might be tackled. Of major importance is the nature and size of the response of physicians trained in medical schools outside Ontario to changes in the average level of physician income in Ontario. Closely related to the question of the general level of response is the question of the distributional consequences of changes in overall supply. Similarly, the size of the potential medical school class in relation to changing income levels is of major interest. An exploration of this issue would likely involve a thorough analysis of the incomes of physicians relative to other professional groups. It could also involve estimates of what income levels for physicians might be if their services were priced in the market place and might involve a consideration of

alternative methods of financing medical education including that of putting the major burden on the medical student himself.¹

The second and third major objectives proposed relate to the allocation of physician activities within the health care sector. The first of these is in the nature of theoretical work on the expected behavior of physicians under different remuneration schemes given alternative assumptions about the principal elements of their objective functions. In addition to profit maximization, either in the guise of maximizing revenues generated or minimizing time spent delivering care maximization, the easy life, the desire to work in sophisticated facilities, as possible elements in the utility function which if included would perhaps yield a different set of expectations about physician behavior under alternative remuneration schemes. It should be emphasized that, the demand-supply model advanced in the early part of this section for examining international or inter-provincial market flow may well prove quite inappropriate. It should also be noted that models are more or less appropriate depending on the particular uses to which they are to be put.²

The third major thrust for research in this area and the second dealing with resource allocation within the health care sector involves the expanded

1. Interesting work on a small part of this problem, namely the sensitivity of medical applicants to changes in income prospects and education costs has been done by F.A. Sloan. See F.A. Sloan, "The Demand for Higher Education: The Case of Medical School Graduates", Journal of Human Resources, Vol. 6, (Fall, 1971) pp. 466-89.

2. Illustrative of this type research is the work by R.G. Evans in building a model of physician behavior that has as a principal component the hypothesis that physicians strive to attain some target income. See R.G. Evans, op. cit., Chapter II.

For a sampling of this type of work down in the United States see M.S. Feldstein, "The Rising Price of Physicians' Services", Review of Economics and Statistics, Vol. 52 (May, 1970); and J.P. Newhouse, "A Model of Physician Pricing", Southern Economic Journal, Vol. 37 (October, 1970).

effort to analyze the behavior of physicians under the different remuneration schemes that currently exist within the province. Both salary and capitation schemes are in existence by special arrangement. The evaluation of physician behavior under these in comparison with control groups chosen from the remaining physician body operating under the fee-for-service scheme could yield a truly exceptional research potential.

The fourth major thrust of research would involve a study of the role played by the colleges and associations of physicians in determining patterns of practice, methods of reimbursement; levels of reimbursement, and so on. This would seem to an especially important area of study in the present environment and would likely increase in value should government move to use its monopsony power in the purchase of health care and in turn should such a move bring about the expected consolidation of the monopoly power of the producer group.

Though we have concentrated our attention in this chapter on the behavior of physicians, similar problems, research methodology and research potential exist with respect to all other categories of health care personnel.

CHAPTER III

Resource Allocation in the Hospital Sector*

I. Introduction

Our concern in this chapter is with resource allocation in the hospital sector and problems towards which economists might direct their efforts in explaining the behaviour of the principal economic actors in this sector. Though we emphasize the role of the "hospital administrator", the roles played by physicians either individually or collectively through their medical staffs or associations, by boards of trustees, by consumers of hospital care, and by government are also considered in part.

The rapid growth of hospital expenditures (which represent some 60 per cent of the health care dollar) over the last two decades has been variously attributed to improved quality, increased use of hospital facilities and rising prices of factor inputs. There has also been the suggestion that, at least in part, waste and inefficiency in the hospital sector could be responsible for cost increases. This view was supported by the Task Force on the Cost of Health Services who wrote:

'In our opinion, inadequacies in hospital management...in the use of expensive, skilled personnel and facilities...contribute in no small way to high costs.'

The consensus that under the present circumstances hospital care is not provided at minimum cost emerges from both theoretical analysis of hospital behaviour given the present institutional arrangements and through observation of hospital

*I would like to acknowledge with appreciation the work of Ms. Kerry Adams in drawing together much of the material on which this Chapter is based and in setting out a first draft of it.

1. Department of National Health & Welfare Task Force Reports on the Costs of Health Services in Canada, Vol. I "Summary"; Vol. II "Hospital Services". Ottawa, Queen's Printer 1970.

behaviour. The most convincing arguments from this body of evidence include the lack of coordination among hospitals which necessitates that each hospital have sufficient capacity to meet peak load demand, the lack of effective community planning, the absence of economic pressures to force the efficient use of capital, and the lack of good cost accounting methods which would contribute to an understanding of the costs of the various hospital services.

In general, it would appear that the barriers to efficient use are inherent in the institutional structure of the hospital system; each hospital operates by and large as a single entity and its actions are not well co-ordinated (if co-ordinated at all) with those of other hospitals. Furthermore, the clinical freedom of physicians likely does much to limit the effectiveness of management in the hospital.

Perhaps the most significant obstacle to efficiency in the hospital, however, is its status as a non-profit institution. The absence of personal property rights, which necessarily implies that the hospital trustees and administrator have no claim to any residual wealth that might be generated, is most likely the prime cause of the lack of incentives to minimize costs. To counteract the "slack" that is likely to accompany the lack of economic incentive to operate efficiently, guidelines and regulations have been set up as substitutes for the incentives that would ordinarily be present in an organization with proprietary ownership. Rules and regulations, however, cannot control all of the vast number of transactions and decisions that must be made in the hospital, and many actions will be taken that rely solely on the discretion of the administrator, trustees, or physicians.

That the competing demands of different physicians for access to beds and ancillary facilities and manpower together with those of patients and members of boards of trustees for particular services do not ultimately result

in the minimization of costs, on average, is difficult to access principally because of the lack of a commonly agreed upon definition of the output of the hospital. This problem is one that derives from two sources: the difficulty in setting out an acceptable measure of output in principle and the comparative lack of data to describe alternative concepts of output.

Resource allocation within, or by the administrator of a given hospital is only one of the decision levels for resource allocation in the hospital sector; indeed, there are three major decision levels that can be considered. There is the overall level of hospital capacity that would be optimal from society's point of view. Secondly, there are the decisions about how one might best employ the hospital facilities that exist within a given region. Thirdly, there are the decisions about how the resources of a given hospital might best be deployed. In this chapter we concentrate on the first and third of the allocation decisions and leave the second to be dealt with in Chapter V on regionalization.

Our tack is first to consider briefly, in Section II, the issues involved in the decisions regarding the overall level of bed capacity in the community. We then proceed to deal in greater detail with the issues involved in resource allocation in a given hospital. In order to understand the decision-making process behind the allocation of scarce resources in the non-profit, non-revenue generating hospital, we need a behavioural model that takes into account the fact that the administrator, patients, trustees, and physicians each influence hospital decisions in a different way. The model must also take into consideration the constraints on the behaviour of these decision-makers and should propose a maximand or an objective function for the hospital which will "adequately" explain the subset of observed behaviour under scrutiny. In Section III, several of the various models of the non-profit hos-

pital that have been proposed are presented, followed by a brief consideration of the effect of the government and the consumer on the constraints facing the hospital decision-maker.

Then in Section IV we present a brief survey of the literature containing proposals for increasing the efficiency of hospitals. In reviewing these proposals one must remember that unless members of the organization have a favourable attitude toward cost reduction, externally imposed pressure for efficiency may result in counter-productive behaviour, such as decreased quality or quantity of hospital services provided. The problem then, appears to be one of devising measures of controlling costs which involve incentives for hospital administrators, patients, trustees and physicians to participate in cost reduction without sacrificing the quality or quantity of care provided.

In the final section we propose an agenda for research by economists in this area.

II The Overall Level of Hospital Capacity

One of the two principal factors involved in decisions about the size and distribution of hospitals required to serve a community is the extent and nature of internal economies and diseconomies of scale in the provision of hospital services from varying sized facilities. This issue has received a significant amount of attention by economists.¹ In contrast, the other factor,

¹For work done on Canadian hospitals, see R.G. Evans, "Behavioural Cost Functions for Hospitals", Canadian Journal of Economics, Vol. 4 (May, 1971), pp. 198-215; R.G. Evans and H.D. Walker, "Information Theory and the Analysis of Hospital Cost Structure," Canadian Journal of Economics, Vol. 5 (August, 1972), pp. 398-418; and R.D. Fraser, Hospital Costs and Efficiency, Special Study No. 13 prepared for the Economic Council of Canada (Ottawa: Information Canada, 1971).

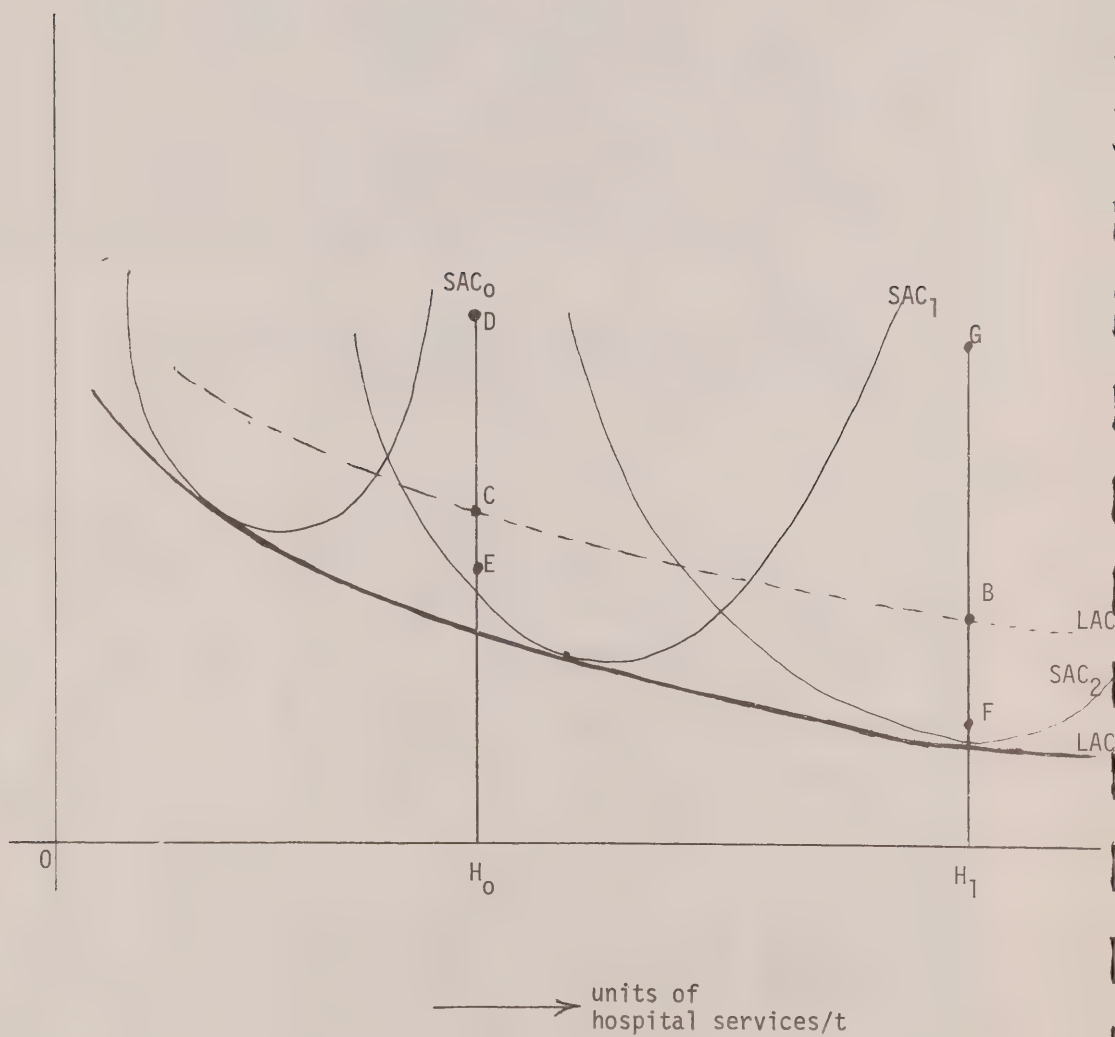
For work done on U.S. hospitals, see J. Lave and L.Lave, "Economic Analysis for Health Service Efficiency: A Review Article," Applied Economics, Vol. 1 (January, 1970).

the direct and indirect costs imposed on the patient primarily, and secondarily on the physician, of the hospital facility being located at different locations relative to that of the patient has received almost no attention. In the present context further work might with advantage be done on the first factor and without question there is significant potential for worthwhile research on the second factor.

With regard to the extent and nature of economies of scale in the provision of hospital care, the present findings that economies of scale are present but are not especially strong are deserving of additional confirmation since all studies to date have been carried out, either with inadequate data on the capital costs of the hospital and its major equipment or in the complete absence of any such data.

Furthermore, the studies carried out would seem to provide an excellent starting point for more detailed case studies. In Diagram III-1 we illustrate the analytical framework of many of the studies on the existence of economies of scale. The ultimate result is the estimation of a cost function such as LAC^* . Depending on the existence and strength of the forces in the hospital sector to minimize costs, LAC^* may lie closer or farther away from the hypothetical, competitive, minimum cost function LAC that might exist if for example the hospital sector was organized in the form of a competitive market. Notwithstanding the actual divergence of LAC^* from LAC , it is the case that the estimation of an average cost function yields a subset of hospitals that are characterized by especially high average costs for a given level of output and also a subset of hospitals characterized by especially low average costs for a given level of output. For example, for all hospitals with an output level of H_0 , the estimated average per unit cost is given by the point C. Average costs represented by points such as D and E actually exist in particular hos-

Diagram III-1. - Unexplained Variations in the Average
Costs of Hospital Care



Source: See text

pitals.

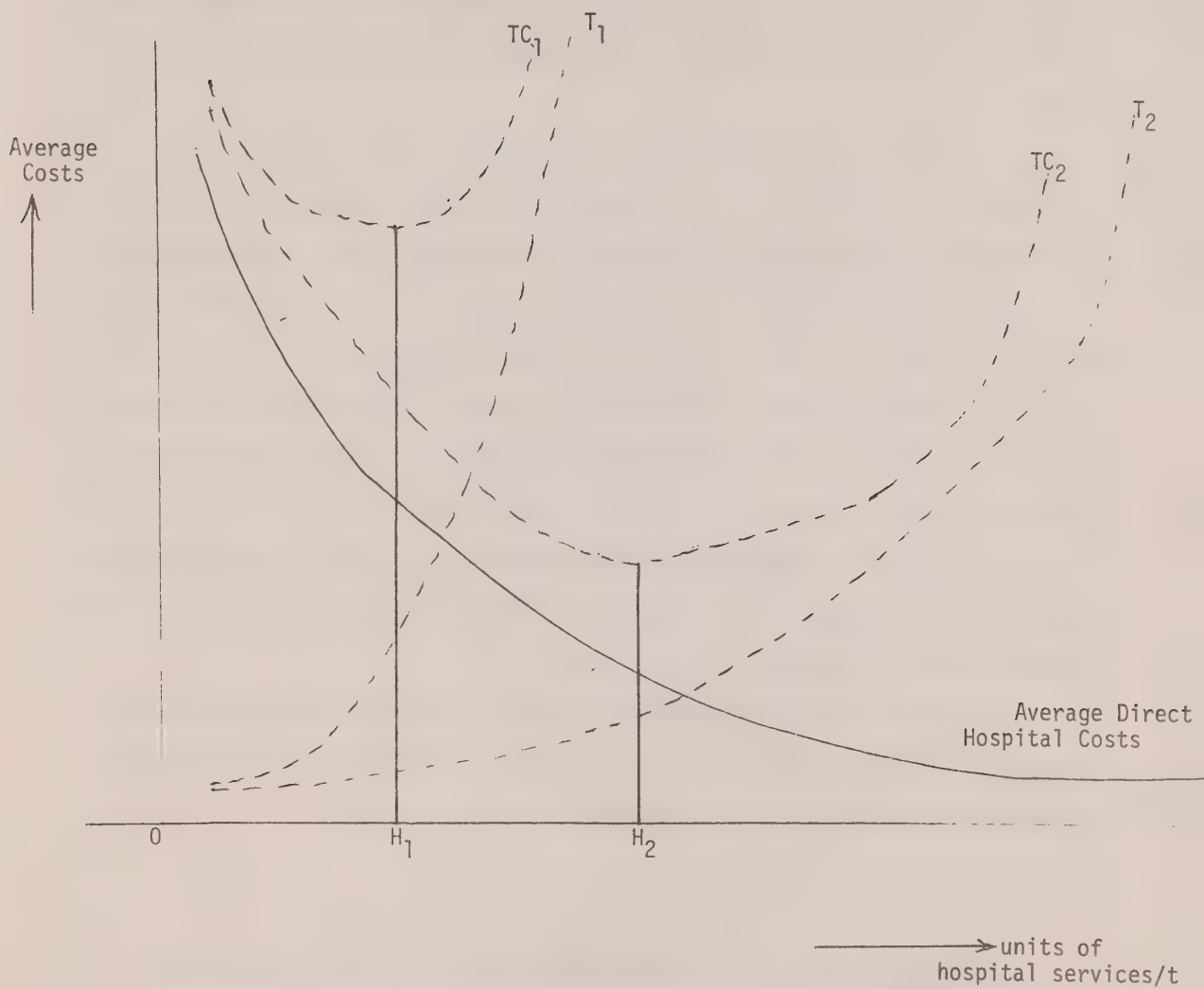
To move the research on the determinants of hospital costs a significant step forward, one promising avenue would be to mount a project that involved a detailed case study of a small subset of hospitals characterized by what appear to be disproportionately high average costs for the given level of output and of a similar subset characterized by especially low per unit costs.

In order to more fully tackle the question of the number and size distribution of hospitals, it is of course essential to estimate the direct and indirect costs imposed on the patient in achieving access to the hospital. In Diagram III-2 we attempt to illustrate the nature of the information sought and its relationship to the direct costs of providing hospital care once the patient is on the premises. The functions T_1 and T_2 are referred to as "transportation" cost functions. They represent the costs of the patient obtaining access to the hospital and include direct transportation costs and the indirect costs associated with reductions in the probability of success in treatment, for example for emergencies, as the patient's residence or place of work become more distant from the hospital location.

In estimating such a transportation cost function for a community, the population density, network of roads, availability of ambulance services, and so on would clearly be of prime importance.

Given estimates of the transportation cost functions, and those of the direct costs of hospital care by size of facility, it should be possible to determine what a "sensible" distribution of hospitals throughout the Province by

Diagram III-2. - The Optimal Size of Hospital



Source: See text.

size and by the nature of the care to be provided. In this research endeavor, the work of economists using the tools of location theory might well prove highly successful. In the same way one can determine a sensible pattern for the location of firehalls, public schools, etc., it should be possible to establish the most sensible location of different types of hospitals by size and kind of care provided. The government might then be better able to make decisions on the appropriateness of the location of proposed additions to the overall Provincial stock of hospital facilities.

III Models of Resource Allocation

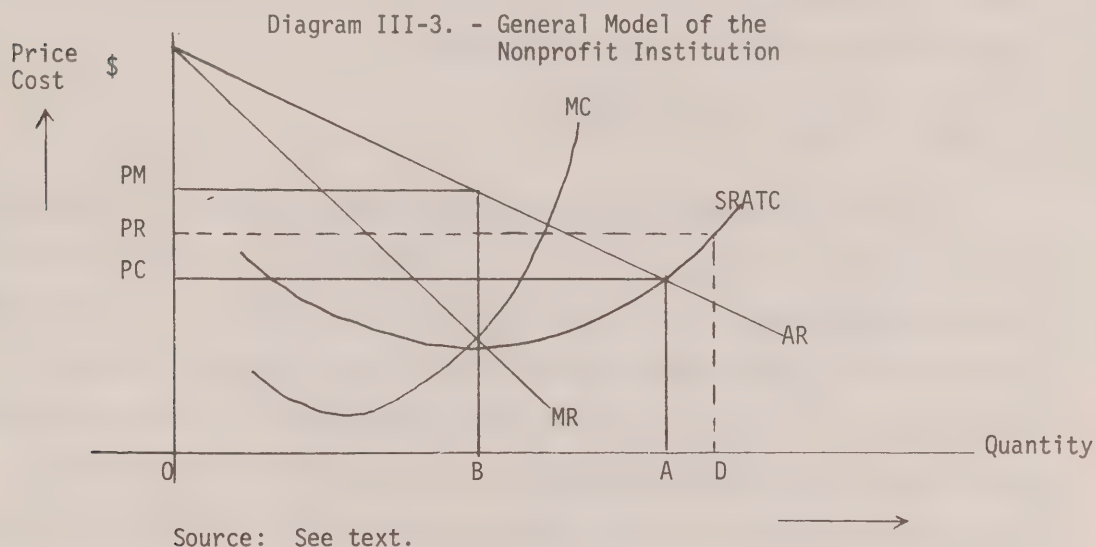
Most of the behavioural models of the hospital that have been constructed may be arbitrarily classified in one of two categories:¹ those based on single or limited performance objectives, for example quantity/quality maximizing models, and those based on more complex specifications of participants in the decision making process. The main issues that these models must deal with are the separation of the decision-making powers between the managerial and medical groups within the hospital and the effect on these decision makers of the lack of economic incentives to minimize costs. All of the models also face the complications arising from the difficulty of defining and measuring quality and quantity of outputs in meaningful ways.

A. Single or Limited Objective Function Models

The major theoretical underpinning of most analyses of hospital behaviour is provided by a general model of the nonprofit but revenue generating, institution. This model predicts that instead of using marginal cost pricing,

1. See Philip Jacobs, "A Survey of Economics Models of Hospitals" in Inquiry, Volume XI, No. 2, June 1974.

the nonprofit institution will produce at the point where price is equal to average total cost, and thus output will be greater than would be the case in the for-profit sector. Diagram III-3 illustrates this model of the nonprofit institution, where SRATC is the short run average total cost curve exclusive of even a "normal" profit; and AR and MR are the conventional average and marginal cost curves of microeconomics.



According to this model, the nonprofit institution will generate output at OA which is sold at the price P_C , whereas output would have only been OB with the price P_M if the industry operated in the private sector by a monopoly firm.

The chief problem with this model in the Canadian setting is that by far the vast majority of Canadian hospitals, though nonprofit institutions, are not in the main revenue generating institutions in the sense that patient consumers directly pay a price for the units of hospital care consumed. Perhaps a more appropriate model applicable in the case of government funding on a per patient or per patient day basis would be one that sees the model in Diagram III-3 revised to indicate an output level that coincided with the equation of average costs to the specified per unit value of reimbursement, for example the output

level D. The determination of P_R then becomes the interesting issue. Another variation of the general model presented in Diagram III-3 is the quantity-maximizing model of hospital behaviour which treats the hospital as if it were much like 'the firm' of traditional microeconomic theory. The quantity-maximizing model assumes that there is only one decision maker, the administrator, and that the quality of care is exogenously given, for example, the "best practice" level of care. By assuming away any variation in quality from "best practice" level, the model implies that the hospital administrator will attempt simply to maximize quantity given his budget constraint to ensure that care is provided at a minimum cost. This model of hospital behaviour is likely to prove inadequate, both because the assumption of an exogenously given quality is untenable and because the complex interaction of the four principal groups of decision makers is ignored.

Since one of the major shortcomings of the quantity-maximizing model is its neglect of the potential variation in the quality of hospital services, it seems worthwhile to consider the construction of a model of hospital behaviour which stresses quality as the maximand in the objective function of the decision maker. Although the quality-maximizing model as usually presented involves the consideration of some level of quality as an endogenous variable, it is naive in that it assumes the administrator is the sole decision maker who bases his decisions on such factors as security, salary, prestige, and professional satisfaction. The physician is only considered to be an input in this model, and he enters the objective function in the same manner as capital equipment does; as a means of providing the best quality of care possible. The quality-maximizing model is thus less useful than it might be because it is too one-sided (it does not give any consideration to the quantity of care provided) and significantly, it does not even attempt to suggest that care is

provided at minimum cost.

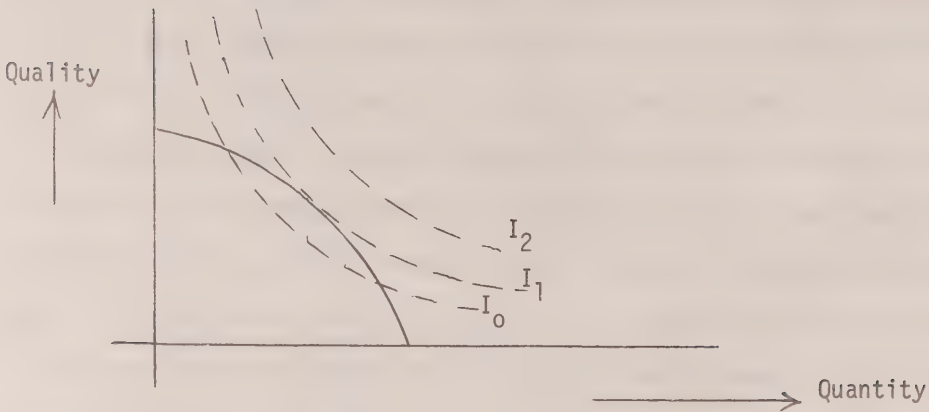
The most frequently used maximand in behavioural models of the hospital combines the maximands of the two previous models and involves a joint maximization of both quality and quantity.¹ This model also grapples with the problem of the management triangle and views the decision maker as an amalgam of trustees, physicians and the administrator, who reach decisions through a bargaining process. The three groups, however, share the mutual objective of maximizing both quantity and quality subject to the budget constraint.

The quantity-quality maximizing model assumes a U-shaped cost curve which shifts up and down respectively as the level of quality is increased and decreased. The problem encountered in attempting to define quality now becomes particularly evident. For example, Newhouse resorts to a "vector of characteristics, not all quantifiable", a nebulous variable at best. Similar problems are encountered in attempting to measure quantity which is usually defined as the number of patient days although it is unlikely that product per patient day is likely to vary with every case.

Despite the practical problems of definition and measurement of the pertinent variables, in theory at least, an average cost curve can be determined which shifts up and down as quality assumes a range of different values. As the shifting AC curve intersects the demand curve at different points, a locus is determined which represents the trade-off curve for quality and quantity, since a different quantity of output is associated with each quality level. Production will occur at the point where the highest indifference curve of the amalgam of decision makers is tangent to the trade-off curve. This situation is presented diagrammatically in Diagram III-4.

1. That is, both quantity and quality are given some weight in allocative decision making. See Newhouse, Joseph, P. "Towards a Theory of Nonprofit Institutions: An Economic Model of a Hospital", AER March 1970, p. 64-74.

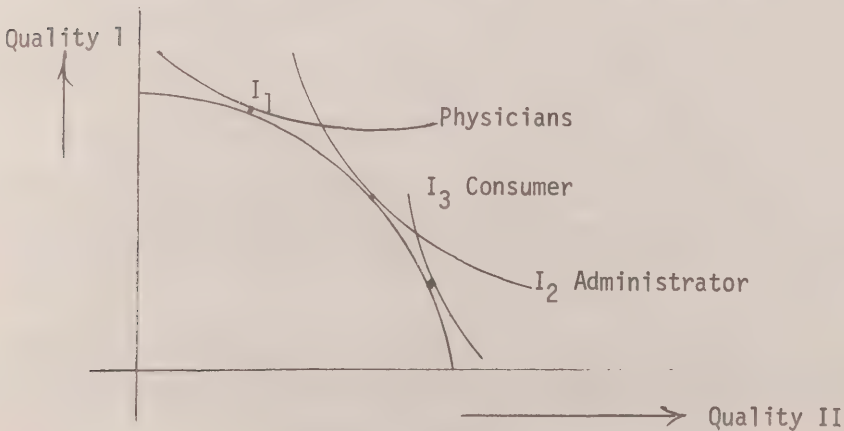
Diagram III-4. - Quantity Maximizing Model



Source: See text.

The level of quality selected is significant not only because it determines the quantity of output provided given the budget constraint, but also because the quality chosen by the amalgam of decision makers (if they can reach a consensus at all) is likely to differ from the level of quality that the "informed" patient consumer would select. In other words, the point on the trade-off curve that is optimal for the "decision maker" is not necessarily the same as that of physicians and patients. The constrained optimum for each of these three groups may well differ and thus it is not clear that the balance between quality levels chosen is socially optimal, as is shown in Diagram III-5.

Diagram III-5. - Differences in Preferences for Quality



Source: See text.

The inclusion of quality in the maximand, though theoretically essential, creates problems because it can be only vaguely defined at best. For example, quality is often confused with comprehensiveness, but it is not necessary for each hospital to have a full range of services within its walls. In short, although quality must be included as an endogenous variable for the model to be complete, its inclusion does add a significantly greater degree of uncertainty to the final outcome.

As a final comment on the quantity-quality maximizing model, we must emphasize that it does not usually deal with the question of whether hospital care is provided at minimum cost.

B. More Complex Utility Maximizing Models

In the second class of models of hospital behaviour, the performance objectives which formed the maximands of the previous set of models are included only as variables in the utility function of the hospital decision maker. While this second approach to analyzing hospital behaviour still makes use of the constrained maximization conceptualization, the objective function in this case is much more complex. In general, for a utility maximizing model to be operationally useful, only a small set of variables relevant to the function can be specified and for these variables any factors affecting their relative costs and utilities must be taken into consideration.

The Trustee/Manager Exchange Model

The trustee/manager exchange model proposed by Clarkson¹ is a basic utility maximizing model derived from observed differences in the behaviour of profit maximizing and nonprofit hospitals. These differences in behaviour

1. Clarkson, Kenneth, W. "Some Implications of Property Rights in Hospital Management". Journal of Law and Economics, Vol. 15(October, 1972), p. 363-384.

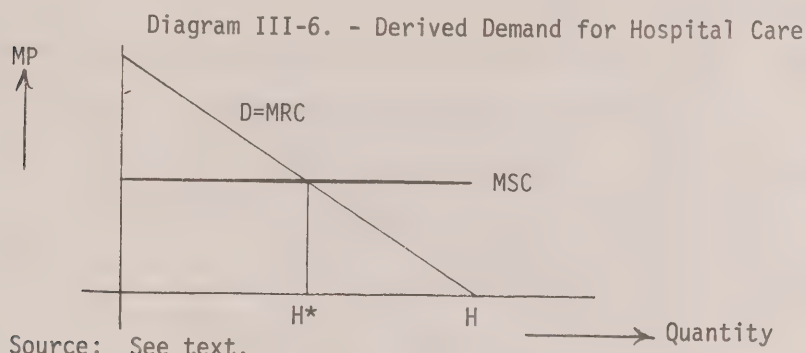
are believed to be largely the result of the difference in property right between the two; namely, that no person in the nonprofit hospitals has the right to any residual wealth that might be generated and no responsibility to shareholders or owners to explain inefficient operations as is the case in the private sector's profit seeking hospitals. The absence of this residual claim in the nonprofit hospital weakens the relationship between a manager's pecuniary wealth and the interests of the organization. Accordingly, the manager may well implement policies which deviate from the profit maximizing rule and which provide him with more nonpecuniary benefits.

The manager's utility function will include the variables quantity and quality. However, they will only be maximized to the extent that his attitudes toward improved staff relations, amenities of the job and so on permit. As a direct consequence of the heterogeneous goals that will enter the manager's utility function when there are no clearly defined property rights, Clarkson hypothesizes that non-profit institutions will have to rely much more on rules and regulations to ensure optimal behaviour on the part of the manager and to prevent too much deviation from low cost production.

It is evident that this model is at best only a partial analysis which considers only the effects of the lack of economic incentives and or responsibility to shareholders or owners on the manager. The value of this and other utility maximizing models is that they strongly emphasize the need for incentives if costs are to be minimized in the nonprofit hospital sector. This approach is more realistic than that of the quantity-maximizing model which assumes that cost minimization will be a natural consequence of managerial decisions. The importance of property rights (or lack of them) as constraints on managerial decisions suggests that there is a need for a market or pseudo-market system of incentives and constraints for publicly funded institutions like the hospital.

Physician Behaviour in the Hospital

With the exception of some of the activities of the outpatients' and emergency departments, almost all hospital utilization is under the direction of physicians. To a great extent, then, hospital output is determined by the aggregate demand of the physicians for hospital factors of production.¹ This organization arrangement may well be inefficient if the physician demands more services than are socially optimal because he considers only the marginal cost to him, ignoring the social marginal cost of producing these services. In Diagram III-6, we illustrate the divergence of private and social marginal costs in the case of one hospital service.



As is indicated in Diagram III-6, if the physician views the marginal cost of an additional unit of service, for example an additional laboratory procedure, as being insignificant or zero, he will request that OH of the service be supplied. However, from a social standpoint there is a positive marginal cost involved in producing the service and thus the optimal level of output is OH* which, depending on the shape of the marginal revenue product curve, may be significantly smaller than OH.

Through their ability to affect resource use within the hospital, the physicians also have the means to significantly affect hospital costs. In

1. R.D. FRASER, "The Economics of Non-Profit Institutions: The Hospital", in Canadian Perspectives in Economics (Toronto: Collier-Macmillan Canada, Ltd., 1972).

recognition of the great impact of physician behaviour on hospital costs and resource use, the Task Force on the Cost of Health Services noted:

"Doctors commit large sums of public money whenever they admit and treat a patient in a hospital. Too frequently, they fail to recognize and accept responsibility for the financial implications of their activities".¹

Pauly and Redisch formulated a utility maximizing model which concentrates on physician behaviour, specifically, it is a physician profit-maximizing model.² The model proposed that the price output decisions of the hospital are designed to maximize the net income of the physicians who have the greatest degree of control over the hospital's operations. The price of hospital services is set such that costs are covered but there remains a residual which goes to the physicians. According to the model, the physicians attempt to maximize this residual subject to the production function and budget constraints.

Although the Pauly and Redisch model attributes a more significant role to the economic motive and places less importance on physical output objectives in the hospitals, it violates the basic assumption that non-profit institutions will set price equal to cover total cost. In the Canadian context this model is not wholly applicable because prices are not determined by the hospitals under the present scheme of government funded health care. It nevertheless represents a starting point for the development of a model applicable to the Canadian scene where the hospital's source of income is different.

1. Task Force Reports, Vol. 2, p. 14.

2. Pauly, M. and Redisch, M. "The Non-For-Profit Hospital as a Physician's Co-operative" AER, LXIII, March 1973. p. 87-99.

Necessary Elements in the Construction of a Model of Resource Allocation in the Canadian Hospital

There can be little doubt that the behaviour of physicians in the hospital setting plays a crucial role in resource allocation in the hospital and thus that this behaviour must play a fundamental role in any model of the hospital that is to be applicable to the Canadian setting.

The effect of third party payments on the constraints facing hospital decision makers is primarily determined by the method selected to reimburse the hospital. For example, retrospective reimbursement on a cost basis does not affect the decision maker's autonomy and accordingly, the output policies of the hospital will remain unchanged. On the other hand, prospective reimbursement even on a cost-basis will impose a greater constraint on hospital (quantity-quality) output decisions.

The method of reimbursement presently employed in Canada varies from province to province although in general hospital budgets are set in advance and the institution is expected to control operations so that the budget constraint is not exceeded. When a deficit is incurred, the hospital is reimbursed on a retrospective basis. Exceeding the budget is not costless, however, because the hospital is likely to find that funds for alternative projects or needs will be restricted as a result of its failure to operate within the pre-determined budget allowance.

In summary, the present system of hospital reimbursement on a prospective budget-allowance basis does impose a degree of constraint on the behaviour of the hospital. Quantity and quality decisions about hospital services provided will necessarily be made with the objective of keeping total costs within the budgeted allowance. The method does not ensure anything more than this adherence to a budget. That is, it does not itself ensure that the

hospital resources are being efficiently employed.

Because the method of reimbursement selected to finance hospitals can influence hospitals output decisions, the government is in a potentially powerful position to control hospital costs via the payment mechanism. Evaluation of the various methods of reimbursement will perhaps lead to the selection of one method most able to help achieve cost control. The difficulty in attempting to improve hospital efficiency is that the added constraint may result in decreased output and/or a deterioration in the quality of hospital services rather than leading to increased efficiency. To date little is known about the efficiency versus quantity-quality trade-off except that it is present to some extent in all hospital behavioural functions. What is clear, however, is that a model of the hospital in the Canadian setting must be based on a more realistic assumption about the way in which the income of the hospital is generated.

As the patient's economic status no longer directly influences his effective demand for hospital services, there is no check on hospital costs from this quarter. In general, the consumer has the illusion that medical care is free and he fails to see the link between the cost of hospital care and his taxes. As a result of this tendency to see only the demand side of the provision of hospital services, the consumer requests only high quality service and guaranteed access in case of need. However, it is to be hoped that with time, greater consumer awareness of the link between hospital costs and taxes will lead both to more public unrest about rising hospital costs and more judicious usage of hospital services. Even if such awareness existed there would still be substantial problems of "moral hazard" since his own taxes will not be appreciably effected by his consumption of care in excess of that that would yield benefits in excess of real production costs. In any

case, the pressures brought to bear on hospital administrators, physicians, and politicians by consumers for the provision of higher quality and greater quantities of hospital care must be considered in the development of a model applicable to the Canadian scene.

Finally, in addition to the incorporation in a model of the Canadian hospital of the roles played by physicians, government, and consumers, the role played by the hospital administrator as the agent of the hospital's Board of Trustees must be considered. In a world of government funded hospital care Boards of Trustees play a much different and probably much more minor role than they once did. They still however possess the legal responsibility for operating hospitals. Thus, the utility function of the "hospital" must capture the principal objectives of trustees, since these objectives will influence directly the actions of the hospital administrator.

In general, the bulk of the work of economists on modeling the behaviour of the "hospital" has concentrated on elements or constraints that are not immediately applicable in the Canadian setting. The methodology employed in these various works is however correct. It involves specifying the objectives of the principal economic actors in the hospital setting and incorporating these in the relevant set of constraints both with respect to the sources of income and also to available medical science and technology including attitudes towards what is thought to be unacceptable quality and what is thought to be "best practice" quality of care. There can be little question that theoretical work by economists of the kind just described could ultimately be of significant value for policy directed towards influencing resource allocation decisions within hospitals.¹ One major caveat should be stated however.

1. In this direction, the continuing work of Philip Jacobs at the University of Concordia in Montreal is worthy of interest.

It is unlikely that a single model of behaviour can be ubiquitously useful. It is more likely that useful models will be directed at subsets of behaviour in the hospital sector. In other words, the researcher would probably be well advised to specify the principal issues about which he is concerned and only then proceed with the development of a behavioural model.

IV Proposals for Achieving Allocational Efficiency in the Hospital Survey

In general, proposals for increasing efficiency in the hospital sector stress the need for better management and for incentives to minimize costs. Three basic approaches to the problem have been suggested in recent literature: comparison measures, incentive reimbursement, and practical methods involving co-ordination of hospitals' behaviour.

Kristein¹

M.M. Kristein advocates the use of comparison measures designed to rank hospitals for efficiency according to eight objective criteria, namely: plant, quality of service, amenity level, range of services, capital investment, teaching, utilization rates and case mix. This proposal rests on the theory that although no system will produce efficiency pressure in the sense that a market system does, internal pressure and incentives may be created through invidious comparison. In other words, hospitals which rate poorly due to high costs will be encouraged to decrease their costs to the cost levels of the average.

Whereas in the USA this method is estimated to represent annual cost

1. Kristein, Marvin M., "A Proposal to Promote Efficiency in Hospitals", Stony Brook Working Papers, No. 114, March 1974.

savings of 10% or one and one-half billion dollars, Evans and Walker¹ estimate that for Canada the savings would be only 5% of hospital costs. Offsetting these cost savings would be the tendency for hospitals whose initial costs are below the group mean to have higher than average cost increases after the comparison.² In other words, hospital costs will tend to converge toward some mean value of average cost and there is no assurance that this mean value will represent efficient production.

Obtaining an objective measure of the inefficiency component of total cost is an Herculean task calling for the "standardization" of all hospitals according to locational differences, hospital size and environmental factors (unions, etc.). Differences in efficiency of production can arise only from two sources; either some hospitals pay more for inputs or they use more inputs to produce a given level of output. In testing for inefficiencies of the latter type, hospitals must be standardized also for age of plant, quality of service, amenity level, range of service, number and duration of teaching programs, utilization rates, and case mix. The result of the model would be a list of hospitals ranked by cost per patient day, as adjusted for the factors listed above. Any differences in relative costs would then be assumed due to inefficiency.

If identification of inefficient hospitals does not produce the necessary pressure on the less efficient to improve or else to cease operations, there is an additional feature of the proposal that is likely to be more effective.

It calls for objective standards to be set by investment review groups so that new funds for replacement and expansion will be allocated only to the

1. Evans and Walker "Information Theory and the Analysis of Hospital Cost Structure", CJE, August 1972, p. 417.

2 One could of course institute a system of rewards for the low cost producers and penalties for high cost producers.

more efficient hospitals. What this does is create a quasi capital market effect which is likely to be much more effective than invidious comparison measures alone.

Lave & Lave¹

The J. Lave and L. Lave model is also founded on the concept of invidious comparison of hospital efficiency and recommends rewarding efficient hospitals and penalizing inefficient ones directly via a reimbursement scheme. In their view, an effective reimbursement scheme will adhere to the following four criteria:

1. it must meet the financial needs of the hospitals.
2. it must be acceptable to administrators and boards.
3. it reimburses on a prospective basis i.e. depends on factors other than incurred costs.
4. it must motivate the hospital to produce high-quality, low-cost curve.

Just as was the case in Kirstein's proposal, hospitals must be "standardized" for such factors as: size, case mix, quality of care, teaching programs, and so on. The scheme necessarily requires a lot of information about each hospital and thus poses the usual problems of definition and measurement of variables.

The model itself takes the form of a function which relates costs to the variables: case-complexity mix, institutional characteristics, utilization, patient characteristics and diagnoses mix. From this function, differential rates of reimbursement among hospitals are determined and hospitals that incur costs that exceed those predicted by the model would be considered inefficient. An advantage of the proposal is that an agreed upon rate of general cost inflation can be built in, although this attempt to control cost

1. Lave & Lave in Silverman, "A Proposal for Incentive Reimbursement for Hospitals". Carnegie-Mellon University, October 1971.

inflation would break down in the face of rapid cost inflation in other sectors of the economy.

Long¹

M.F. Long deals with only one aspect of the problem of inefficiency in hospitals; namely, that peak loads and fluctuations in demand over time impose added capital expense in the form of increased investment in plant and equipment. At the root of the problem is the hospital's apparent disregard for the rate of return to the community when it invests funds, and this results in overbedding and duplication of facilities. Long believes that an optimal level of required equipment can be obtained for each hospital and that this level would be sufficient to ensure adequate services to meet demand given that hospitals co-ordinate their actions and develop external as well as internal flexibility.

There is good potential for external flexibility as peak loads seldom occur in all hospitals simultaneously; thus fewer beds are required overall if hospitals co-operate and send patients to other hospitals when they are full. There are limits to this external flexibility defined by the costs of moving to an integrated system in terms of physician time wasted as patient visits become spread out between hospitals.

Fewer beds will also be required if the hospital develops greater internal flexibility by using "swing" beds which can be switched among departments as needed. Along with the use of swing beds, personnel scheduling systems can be designed to take advantage of knowledge of anticipated fluctuations. Internal flexibility has an upper limit determined by the probability that a department will have demands on its services in excess of the planned amount. As the number of swing beds (required to maintain the same level of protection

1. Long, Millard F., "Efficient Use of Hospitals", Economics of Health and Medical Care, Ann Arbor, University of Michigan.

as fixed beds) rises, a point will eventually be reached where savings from further purchase of swing beds would be offset by added expense.

Long has successfully pointed out that hospitals must be forced to recognize the opportunity cost of capital and to invest less in facilities that are duplicated elsewhere. However, he has given only techniques to decrease costs rather than incentives to do so.

Griffith, Hancock and Munson¹

The solution to hospital cost inflation proposed by Griffith, Hancock and Munson also emphasizes practical ways to contain costs rather than incentives to do so. However the proposal is of interest because it recognizes that an essential ingredient of cost containment is the determination of measures by which trustees, administrators and physicians can gauge their performance according to norms and standards. Their proposal also has the advantage of being compatible with the present system. That is, its implementation is reasonable within the structure of the hospital industry.

The proposal advocates cost controlling activities in four main areas: facilities planning on a community level, scheduling of admissions for each hospital, control of facilities utilization and the quality of service of the medical staff, and administrative control of manpower and expenditures.

First, they call for effective planning and control of physical resources by co-ordinating the actions of hospitals to eliminate overbedding and duplication of expensive equipment, and to encourage sharing of laundry, computer and laboratory services. Given the hospital's capacity, admissions scheduling would be designated with the objective of maintaining a 90% occupancy factor so that emergency patients would not be endangered. Facilities utiliza-

1 . Griffith, Hancock and Munson, "Practical Ways to Contain Hospital Costs", Harvard Business Review, Nov.-Dec. 1973.

tion is almost entirely controlled by the doctors; thus they must be encouraged to contribute to cost reduction by not creating unnecessary demand and doing diagnostic work prior to, or on the day of admission for elective patients. Finally, the authors recognize the need for engineering studies to design a budget-control system which encourages sound management of expenditures on a day-to-day basis. This would require the adoption of a performance standard and frequent comparison of achievement against it.

It is evident that this is the proposal of businessmen and that it relies on the premise that the hospitals would be much more efficient if such traditional business practices as planning and management control were employed. However, the proposal does not suggest any economic incentives for the hospital industry nor does it explain how a favourable attitude toward cost reduction is to be developed in the hospital sector. Effective management techniques will only be rigorously applied after administrators, trustees and physicians are strongly enough motivated to accept the necessary changes and hardships involved in cost reduction.

Walker¹

The work of H.D. Walker in this area though still in progress is an example of the work that needs to be done. Walker is attempting to set out a procedure for estimating the impact of case mix on hospital cost with the ultimate goal of setting out the parameters of a formula for use in reimbursing hospitals for the value of the services they rendered as estimated by the formula.

1 . H.D. Walker, "Reimbursement of Hospitals Based Upon Case-Mix", paper presented to the Symposium of Canadian Health Care Economists, Queen's University, September 5 and 6, 1974.

Stoddard¹

Of equal interest as Walker's work and in a similar state of completion, namely being "in progress", is the work of G. Stoddard. He is attempting to operationally define the content (and limits to it), of the treatment package available for an "episode" of illness. If for a number of illnesses the treatment package for the entire episode, appropriate from the standpoint of "best practice" care, can be defined within fairly narrow limits then it might be possible and advantageous to establish a reimbursement scheme that was a function of the number of episodes of various kinds handled. The hospital would then be given the freedom to determine what it deemed to be the most efficient means of providing the medically prescribed treatment package. For example, it could make more satisfactory decisions on whether care should be on an in-patient or outpatient basis than is possible under financing schemes that reimburse hospitals on the basis of the number of outpatients, the number of in-patients, and so on.

Finally, we might note some of the basic work on disease costing² that has been going on for sometime and to which Stoddard's work is related. The objective of this work is to determine the actual value of the services rendered to patients with different diagnoses. In particular, attempts are made to estimate the value of the resources that are in the nature of central services available in the hospital including both physical and human capital. Of interest in this general area of research is the question of whether the costs of the system that would ultimately have to be mounted to determine on an ongoing basis the real value of the services rendered to particular patients would be exceeded by the benefits to the hospital or health care delivery system

1 . G. Stoddard, "An Episodic Approach to the Demand for Health Care", Dissertation Prospectus, Department of Economics, University of British Columbia, December 1974.

2. J. Babson, Disease Costing (Manchester: Manchester University Press 1963).

of having such information available. In this respect the decisions of profit making hospitals in the United States on the extent to which they decide to gather detailed cost data in order to set prices to patients would be of not insignificant value.

V Research Agenda

It is evident that the success attained in the research and development of economic incentives in the hospital industry will be a critical factor in the speed and extent of implementation of programs designed to help reduce hospital costs. It must be determined whether it is possible to create worthwhile incentives within the present system of government funded health care which is based on fee-for-service reimbursement to doctors and retrospective reimbursement to hospitals for incurred costs of services provided.

What is clear is that simple imposition of regulations to secure "cost reduction" will be ineffective and potentially detrimental to the quality and quantity of health care provided in this country unless the hospital administrators, trustees and physicians are willing to make the necessary sacrifice and adjustments for cost control.

This seems to demonstrate the need for work on the expected behaviour of the principal economic agents of the hospital in order that the impact of any particular incentive scheme can be evaluated. Such work on the behaviour of the "hospital" would seem to constitute an appropriate prime research target for economists.

Secondly, there is work on the nature of the treatment packages provided by the hospital that is required if reimbursement plans related to the number of therapeutic episodes are to be evaluated, let alone developed. Even if the reimbursement schemes remain much as they are today, however, it may well be of interest to society to develop a monitoring system that would seek

information on the nature of the treatment package provided in different hospital settings to patients with specified diagnoses.

A third major thrust for research involves a consideration of the overall stock of hospital facilities and its distribution in relation to the population. Research efforts by economists would seem to be particularly well suited to address the questions of: the extent and magnitude of economies of scale in the provision of hospital services, the nature of the "transportation" cost function, and then the related issue of the sensible location of hospitals by size and type of care (including active treatment, chronic, convalescent, nursing home care, etc.) throughout the Province. Of course, this methodology would also be appropriately applied to the network of ambulance services available in a community or throughout the provinces, the location of laboratories, and so on.¹ Study of the experience of the Scandinavian and Eastern European countries might well prove advantageous. For example the USSR has been involved in planning the use and the development of its hospital system for some years.

Before concluding our discussion of resource allocation in the hospital sector it is worth emphasizing that data problems have in the past proved effective barriers to many interesting avenues of research. There appears to be light on the horizon however. In particular, the Hospital Medical Records Institute has been developing its data bank rapidly so that it will soon cover the bulk of Ontario Hospitals. Data available, at least in principal, will include both financial data and data of the number and nature of the services rendered to individual patients as related to the different diagnoses that requires access to historical data is of course not much helped by this new data assembling resource.

1. Of particular interest in this regard is the work of R.A. Taylor. See R.A. Taylor, "Hospital Utilization and Ideal Output: Theory and Evidence", Working Paper #17, (Department of Economics, Northern Illinois University, DeKalb, Illinois).

Chapter IV

Consumer Behavior in the Health Care Sector

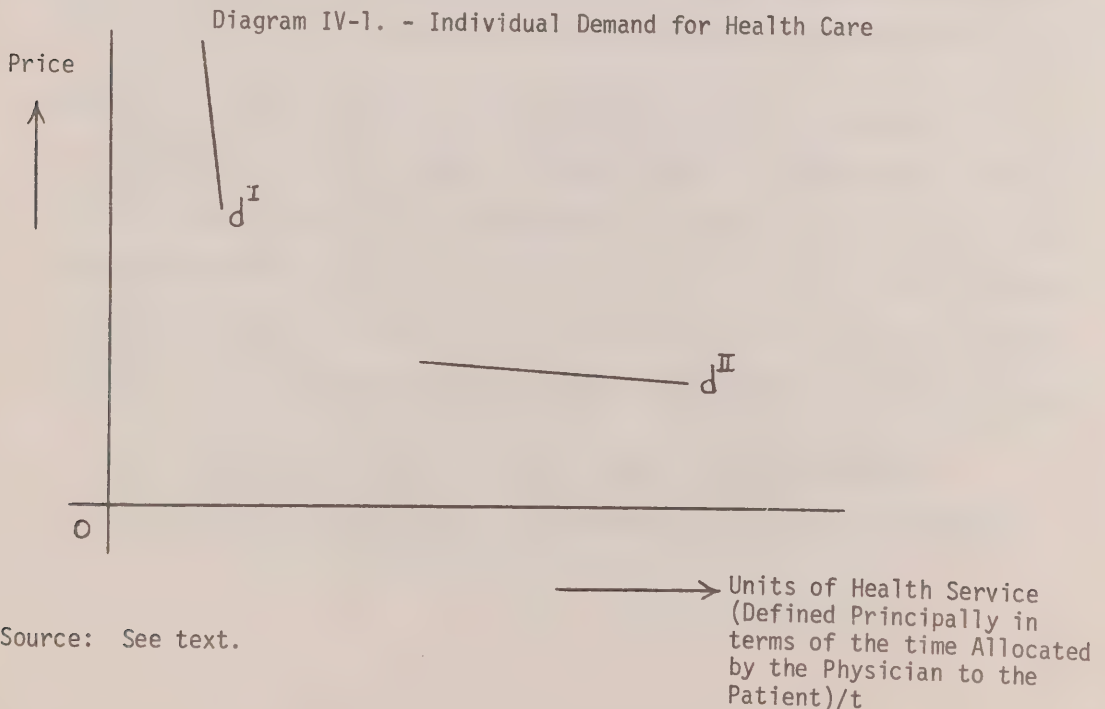
I. Introduction:

The behavior of consumers of health care is of course germane to almost every economic problem in the health sector. The justification for not giving the consumer a more prominent role as a principal economic agent in the sector than we have or as might be expected in line with studies of other markets stems from a number of sources. First and foremost is the nature of health care itself and the usually presumed limited ability of the consumer to evaluate the health care required and or provided. As a result it is often argued that the distinction between demand and supply cannot generally be made. Secondly, the present organization of the health sector does not permit the consumer to act out his more traditional role in the major resource allocation decisions in this sector. And thirdly, it is widely assumed that the political climate in Ontario is such as to all but preclude the possibility of making changes in the organization of the sector that would give the consumer a more active role to play.

Depending on one's judgement of the validity of the last argument further research effort directed towards an analysis of consumer behavior will be more or less valuable. However, even if one judges to be low the potential for significantly increasing the importance of consumer decision making, it can still be argued that his behavior is of enough importance under the current arrangements to make the continued study of it worthwhile.

One aspect of this behavior that is especially important is the analysis of the factors that influence an individual's decision to seek

care from the health care delivery system. The demand function of an individual for care is not easy to construct. In addition to those factors thought generally to affect demand, there are ones that appear to play a special role in the health sector. They include: the choice of the time period over which the function is specified, the role played by the physician, the all or nothing nature of the devices, option demand, and so on. In spite of these problems, as a first approximation in Diagram IV-1 we illustrate the nature of an individual's demand for health services by the demand functions d^I and d^{II} . The first of these is thought to be relatively price inelastic, which would correspond to the demand for those health services that the consumer perceives will be life saving services or ones that will relieve, forestall or prevent intense pain, suffering or long term disability. The second curve lying to the right of the first is thought to be more price elastic and to correspond to the demand for health services that the consumer perceives to be partial substitutes for other goods and services that might be purchased; for example, in-patient hospital care for home care by the family unit, or psychotherapy sessions in mid-February instead of a ten day trip to the Caribbean.



Clearly all individuals do not possess the same demand curves for health care as shown by the functions d^I and d^{II} . Depending on social-economic status, and on several other factors the position of the demand curves may well vary significantly among individuals. This general issue of the socio-economic, geographic determinants of demand and the way in which such factors influence the distribution of health care services in the province is taken up separately in Chapter IX¹. In the present Chapter, we largely ignore these factors and instead concentrate our discussion on the "average" consumer of health care.

In pursuing our review of research potential in this area of consumer behavior we first consider selected aspects of the demand of the average consumer for health care including the divergence of private from social demand and the interaction of consumer and physician demand. We then proceed to consider the potential for establishing some variant of the "tiers garant" class of government funding mechanisms; these schemes usually involve some element of co-insurance. Finally, we set out our suggestions of research priorities.

¹ In particular, much of the literature on the nature of consumer demand and its determinants that is discussed in Chapter IX is relevant for the issues raised in this Chapter. Indeed this Chapter and Chapter IX might well be taken together.

II. The Nature of Health Care and Consumer Demand

On the assumption that there is a significantly non-zero probability of being able to alter the organization of the health delivery system in Ontario in the direction of involving the consumer directly in the funding of health care, there are a number of gaps in our knowledge of the nature of consumer demand that require the attention of researcher. These are gaps which if filled would likely permit the adoption of the better of several possible methods of directly involving the consumer in the funding of health care.

A. Private versus Social Demand

One of the more frequently advanced arguments by proponents of government funded health care is that a significant part of health care is in the nature of a public good rather than a private good. One of the most commonly raised examples was the health care activity of clearing a malaria swamp. If one individual in a malaria infested community deemed it worthwhile to expend the resources necessary to clear the source of the malaria not only that individual but the entire community was the receipt of the benefits stemming from the one individual's actions.

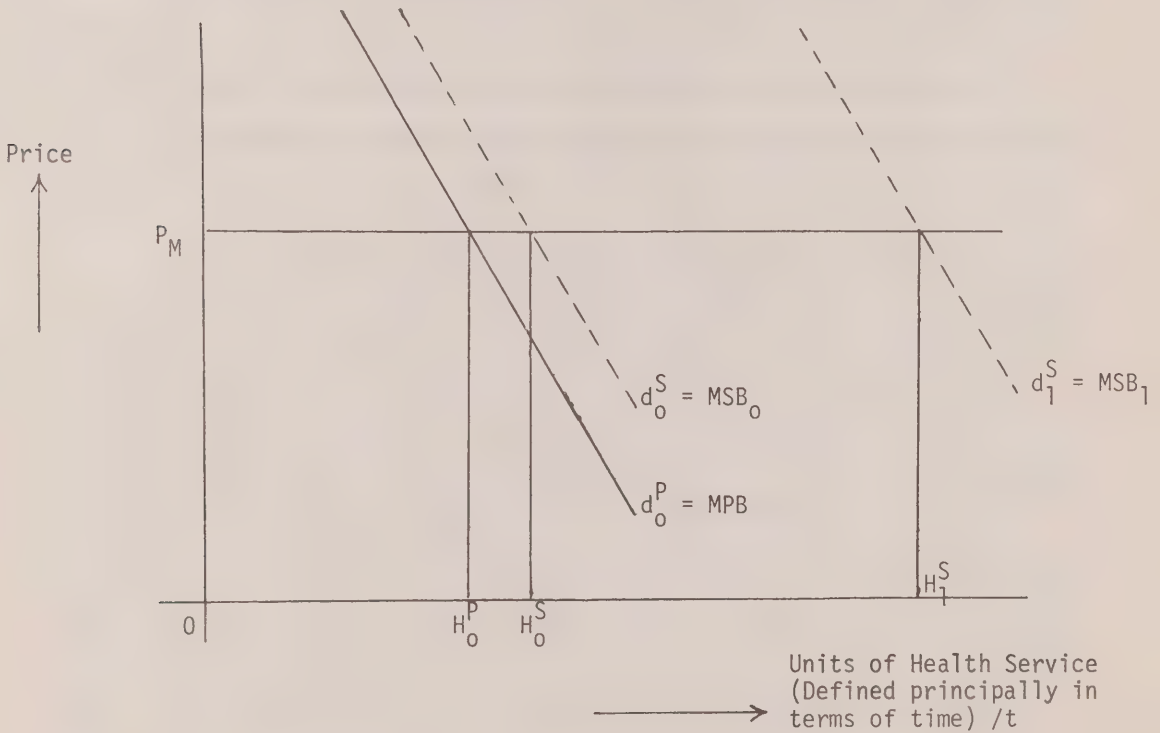
In the general case, than, with individual's making resource allocations decisions with only their own private costs and benefits in mind, the total social benefit of an activity would be grossly underestimated and thus fewer resources would be directed towards provision of the health service in question.¹ The argument then continues that government decision

¹ J.M. Buchanan & M.Z. Kafoglís have demonstrated that over provision of resources is also possible. See J.M. Buchanan and M.Z. Kafoglís, "A Note on Public Goods Supply", American Economic Review, Vol. 53 (June, 1953), pp. 403-414.

making through the funding mechanism would bring society closer to the optimal allocation of resources.

In Diagram IV-2 we have illustrated this issue. The demand function d_0^P represents the benefits of the health service as viewed by the individual. The demand functions, d_0^S and d_1^S , in turn represent the total value of the

Diagram IV-2. - Private versus Social Benefits of Health Care



benefits as seen by society. The first of these, d_0^S , illustrates the case in which the net positive externalities or spillovers associated with one individual's consumption of health services are relatively small. In this case, if the price P_m represented real social marginal costs, the output

level of H_0^P would be only marginally less than that which is presumed to occur with government funding, namely H_0^S .

On the other hand, if the externalities are significant the social demand function might take up the position of d_1^S or indeed a position much farther to the right of d_1^S . In this case, the output level that derives from the decision making of individuals in a market price system would be significantly different from the socially optimal level of output, H_1^S .

In actual practice, health goods and services are likely to form a continuum from those that are analogous to the clearing of the malarai swamp at the one extreme to those at the other extreme that resemble what must be close to the pure private goods, for example the removal of an ingrown toe nail. Now whether the individual consumer's choice backed by purchasing power leads in general to significant resource misallocation is an open question, the answer to which depends on extent to which health goods and services fall to the one or the other end of this continuum. Casual empiricism suggests that in most well developed countries with the near complete absence of communicable diseases the largest bulk of health services are indeed personal, private goods with little if any externalities. Those health goods and services that fall outside this general description could likely be effectively provided through specific arrangements for each exception.¹

¹ In a recent study of the extent of expenditures on non-personal public health goods and services, (that does not precisely describe public goods on the manner that we have but yet provides what are likely to be fairly indicative figures), expenditures on this class of care were seen to vary in eleven countries from .3 to 2.8 per cent of all expenditures on health care; for Canada the figure of 1.0 per cent was estimated. See E. Abel-Smith, *An International Study of Health Expenditures*, Public Health Papers, No. 32 (Geneva: World Health Organization, 1967) pp. 56-60.

One facet of the problem of the extent of externalities relates to what might be referred to as "quasi-externalities". These are defined in terms of the resource impacts that one government (welfare) programme has on another. In the health care case, it might well be that consumers if left to their own devices would purchase a smaller amount of health care at one period in life and in a subsequent period might find themselves the recipients of benefits under a welfare scheme for example which would not have been the case had they consumed a larger amount of health care in the earlier period. Again casual empiricism suggests that these quasi-externalities are relatively small; at the same time however, it appears to be the case that they are growing.

Of course whether government intervention into the health care area would lead to optimal output from society's point of view is itself dependent on many factors including the costs of acquiring information on the just what is optimal output, on the nature of the government intervention, and so. Interestingly, with most insurance type schemes that do not interfere with decisions of the consumer to avail himself of the resources of the health sector, optimal output would only be achieved coincidentally; that is, it would only be achieved if the quantity of care demanded by the individual at a zero or .1 P_M price¹ equalled optimal output, or the level H_1^s in Diagram IV - 2.

"Harder" information would seem to be required to better evaluate these aspects of the nature of health care and the consumers demand for it

¹ This price is descriptive of the present Ontario scheme that reimburses the physician at a level of 90 per cent of the approved fee-for-service schedule and permits the consumer to be billed directly for the remaining 10 per cent.

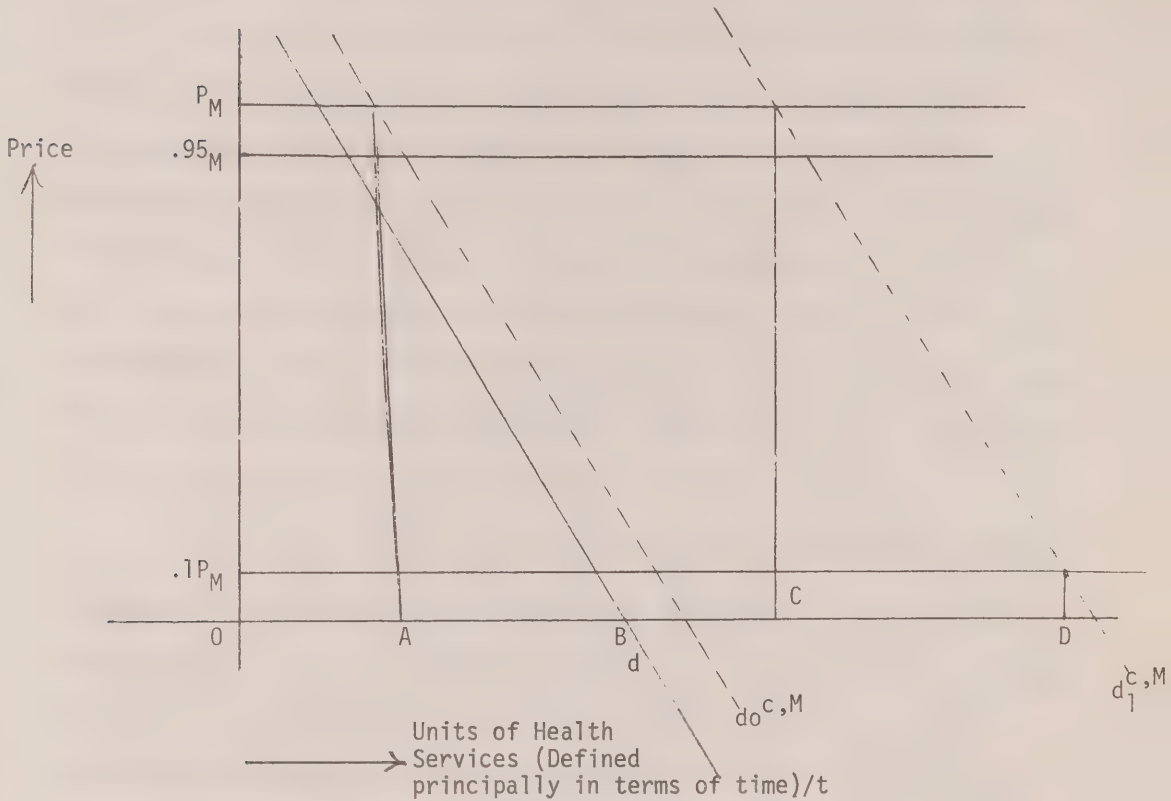
in order that better judgements might be made of the comparative effectiveness of alternative health care financing schemes.

B. Interaction of Consumer and Physician Demand

Perhaps the overwhelming argument in favor of government funded health care turns on the believed inability of the consumer to satisfactorily judge the value of health services to himself let alone any spillovers there might be to other individuals. In turn, the consumer is led to place an especially strong, if not complete, reliance on the work of his physician. Because of this inability on the part of consumers, not only is health care insurance justified but also licencing procedures and the accompany educational requirements and body of regulations.

In Diagram IV-3 we attempt to illustrate this argument. The demand function d^C represents the value-quantity relationship as perceived by the individual. The functions $d_0^{C,M}$ and $d_1^{C,M}$ in turn represent two of an infinite number of alternative value-quantity relationships for the individual after he has taken the advice of his physician. The first of these $d_0^{C,M}$ in relation to the individual's own demand function d^C portrays the interaction between consumer and physician demand for that subset of health goods and services for which the consumer is a relatively good judge of the expected value of the goods and services. In contrast, the second value-quantity relationship $d_1^{C,M}$ in relation to d^C would portray that subset of health goods and services for which the consumer relied heavily on his physician for advice. The first visit of the consumer to the physician is an example of the first subset whereas many complex surgical procedures that

Diagram IV-3. - Interaction of Consumer, Physician and Social Demands for Health Care



are carried out in hospital would be examples of the second subset.

Again, over the entire set of health goods and services there is likely to be a continuum of different sized divergences between the individual's own perceived demand for the care and his demand modified by the advice of his physician. If the bulk of health goods and services were of the first type with consumer sovereignty a meaningful characteristic, the arguments in favor of government funded care for this reason would by definition be unsupported. Clearly it is presumption that the bulk of health

care is of the second type, characterized by the almost total lack of consumer abilities to evaluate care, that provides the basis of the argument in favor of government funded health care.

The essential factor in this matter is the expectation that the physician will be led to recommend more extensive and expensive treatment packages than are in the interests of the consumer and that would be seen to be so if the consumer had the requisit amount of knowledge about the nature of health care or had access to persons who could advise him. With government funded health care the extent of over use and thus resource misallocation is of course further enhanced by the removal of the price barrier to the patient and the requirement that the consumer can back his demand with dollars as far as the physician is concerned.

Contrary to the casual empiricism of the proponents of government funded health care that consumers are all but ignorant, the casual empiricism of many physicians, but especially general practitioners, is that significant portions of their patient load should not be "wasting their time". This statement could of course be little more than the physician's interpretation of the fact that with a zero price at the highest a price of 10 per cent of the fee, the quantity of care demanded by the consumer will yield a level of consumption at which the value of the care to the consumer is but a small fraction of the real social costs of providing it.

Casual empiricism aside, it has been shown¹ that adult patients direct the largest number of their health care demands to persons outside

¹ K. White, T.F. Williamson and B.C. Greenberg, "The Ecology of Medical Care", New England Journal of Medicine, Vol. 265, (1961), p. 885.

the health care sector as traditionally defined. From some 1000 adults over the course of a month, White and his associates have estimated that there are 750 demands for health care. Of these only 250 are made of physicians. The remainder are made of a wide variety of persons including family members, retired nurses (public health or otherwise), priests, ministers, and so on. The many decisions made by consumers in the present context or in the pre-government insurance era to approach a specialist rather than a general practitioner provide further evidence of consumer decision making. Clearly then the consumer does exercise a great deal of sovereign decision making among alternative avenues to the provision of care.

One other feature of the health care delivery system that plays an essential role in determining the success of consumer decision making and that in practice tends to thwart the consumer is the collusive way in which the medical fraternity appears to husband information about differential abilities of its members. There are many other industries that produce products or services about which the consumer can be said to be initially all but completely ignorant and therefore about which the consumer must have recourse to the producers and, or, to other sources for information about the nature of the commodity. In other words, the complexity of the good itself and the ignorance of the consumer in an unadvised or uninformed state is not in itself sufficient reason to view with alarm the outcome of the market price system unless there are constraints imposed on the consumer in his search for what he perceives to be required information about the expected abilities of the physician from whom he intends to seek

care.

In general, a greater understanding of the interrelationship between consumer demand and the role of the principal supplier is required if different financing schemes are to be reasonably evaluated since the nature and extent of the interrelationship likely varies widely over the continuum of health goods and services.

What we should emphasize is that the change from a market price system to a government funded system may well do little if anything to alleviate the problems discussed in this section. It may simply shift the responsibility for dealing with them from the consumer to the government. As consumer, the individual must still place heavy reliance on the recommendations of his physician. He no longer under government financed health care need bear the financial costs of following the physicians recommendations of an expensive treatment programme. Unless government is prepared to intervene in the diagnosis and prescription part of the physician's activities, the consequences of the patient-physician decision will at the minimum be no less wasteful of resources than it was previously but in general will be much more so.

III. Alternative Methods of Directly Involving the Consumer in the Funding of Health Care

As noted in Chapter II, there are several major variants of the "tiers garant" class of government funded health care. All of these variants involve the direct billing of the consumer by the physician for services rendered and the majority of these involve some degree of direct consumer contribution to the payment of the physician for services rendered; that is, some degree of co-insurance. In this section we plan to describe briefly the principle variants of "tiers garant" funding and discuss the major problems of the different variants. In our discussion we consider as necessary characteristics that the funding system should not pose a barrier to the consumption by an individual of needed health care, should protect the individual against catastrophic health care expenditures, should provide incentives for the efficient use of resources, and should be unordinately costly to administer.¹

A. Alternative Variants of "Tiers garant" Government Funding

Full Government Reimbursement:

At one extreme is full government reimbursement of the consumer for the total amount originally billed by the physician. This variant usually involves an agreement between government and the medical profession

¹ To these criteria, M.S. Feldstein has added two further ones, namely that any scheme should avoid a large tax increase on implementation and should be generally acceptable. See M.S. Feldstein, "A New Approach to National Health Insurance", The Public Interest, Number 23 (Spring, 1971), pp. 93-94.

on the level of the fees at which billing and reimbursement will take place. The major advantage of this variant is that the consumer is made clearly aware of the value of resources that the physician claims were expended on the consumer's behalf. In turn, three checks on resource use are thereby made possible. First, to the extent that the consumer will modify his use of health care system if he is fully aware of the costs incurred on his behalf, resource might well be reduced with this information available. Secondly, as a taxpayer the consumer can be expected to serve as a monitor, however weak, of any tendencies on the part of the physician to bill for services not rendered. Thirdly the government under this scheme possesses fairly detailed information on the services being rendered by physicians and is thus in a position to evaluate the patterns of health care delivery in the system. These implications of consumer and government awareness of resource use in the system is characteristic of all variants of the "tiers garant".

Annual Deductibles

In order to put greater force into the consumer check through his budget constraint on the extent of resources expended towards his health care, it is necessary to involve the consumer in making a direct contribution out of his disposable income in paying the physician. The remaining variants of "tiers garant" funding have this involvement of the consumer. The first of these variants involves an annual deductible. The feature of these variants is that the consumer pays for all services rendered to him

up to a certain level. This level is set a number of ways, of which the two principal ones are the setting of a fixed annual deductible that applies universally and the setting of fixed annual deductibles whose level varies with the income of the individual or household unit. The last possibility could of course involve a zero deductible for all consumers whose income fell below some arbitrary level, for example, incomes that led to zero income tax. In this way, it is hoped that any undesired distributional consequences can be mitigated.

The principle thrust of the annual deductible would appear to be that of providing insurance against disproportionately large health care expenditures in any one year.¹ In contrast, the bite imposed by the consumer's budget constraint on the unjustified use of health care resources is directly related to the level of the annual deductible. If set sufficiently high, significant bite could be expected to exist.

Charges per Service, per Case or per Episode

The third major class of variants of "tiers garant" funding involves the patient in paying some portion of the costs of each service or each treatment package (for defined cases or episodes of illness) rendered to him. This portion to be paid can be a fixed dollar amount per service or a fixed percentage of the total value of each service. In turn, either the level of the fixed dollar amount or the particular percentage levied can themselves be varied as determined by the consumer's level of income. In the

¹ The "Major Risk Insurance" scheme advocated by Feldstein for the United States is an example of this annual deductible scheme. See Feldstein, op. cit. pp. 99-105.

extreme, these fixed dollar amounts or the percentage could be zero for all consumers whose income fell below an arbitrarily chosen level, again for example, incomes that led to zero income taxes. In this way the distributional consequences of charges on consumers can be all but eliminated.

The principle thrust of this variant of "tiers garant" would appear to be aimed at providing the check of the consumers budget constraint on each service purchased. Insurance protection against relatively high health care expenditures in a given year is not provided for by this variant.

Combination of Charges per Service and Annual Deductibles

Finally, there is a wide variety of combined systems that involve charges per service as well as an annual deductible and may at the same time involve these charges and deductibles being geared to the income of the patient. With these schemes it is hoped that insurance against major catastrophe can be provided while at the same time the constraint imposed by the consumer's budget can be brought to bear on each decision to purchase health care.

B. Objectives and Problems of Co-Insurance

The principal argument against the use of co-insurance schemes relates to the presumed deterrent to the consumption of health care thereby instituted. In the foregoing description of some of the principal variants of "tiers garant" or principally co-insurance, we have noted the ways in which in principle one can avoid deterrents to consumption that are felt to have undesired distributional consequences. In general, however, we postpone our discussion of the distributional impact of government funded care to Chapter IX.

In this section we wish to concentrate on the potential for achieving the appropriate level of deterrence on the assumption that the lack of any financial deterrent in the current situation entails a significant misallocation of resources. If we refer to Diagram II-2, again, the introduction of government funded health care which lacks a deterrent fee will lead to consumption far in excess of the social optimum unless the supply of services poses a constraint.

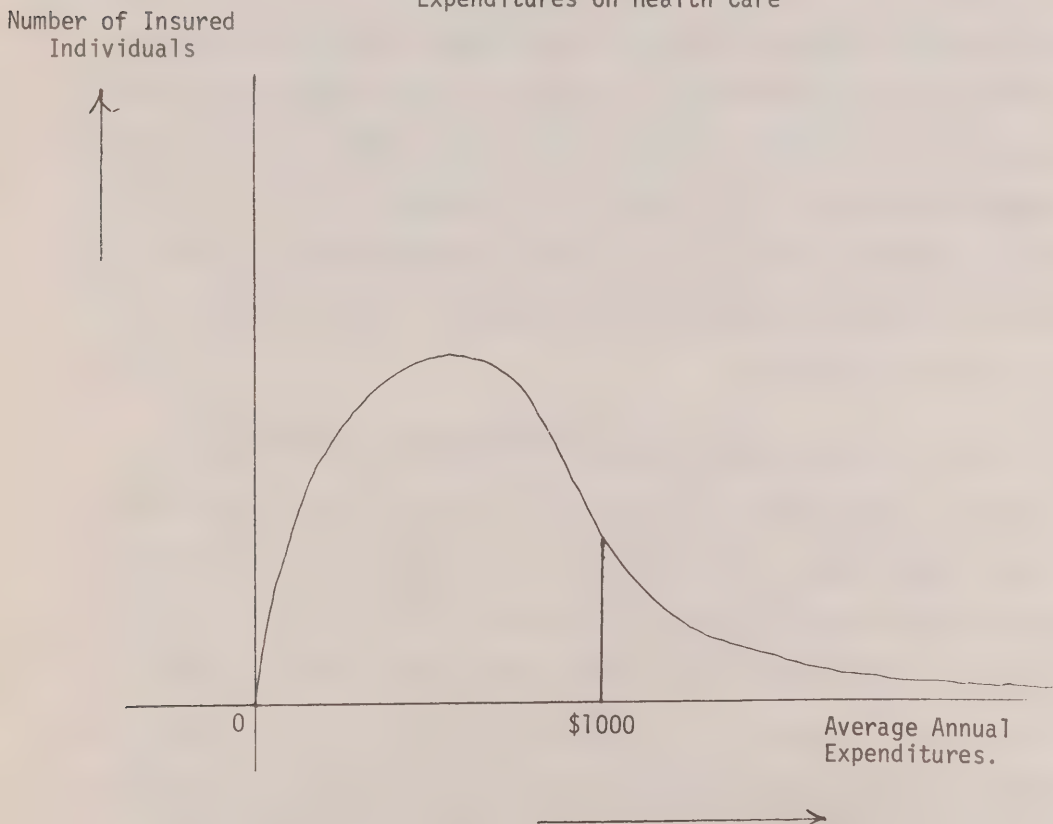
If of course health care were in the main characterized by significant net positive externalities as discussed with the aid of Diagram IV-2, than the absence of a deterrent fee might well allow for a closer approximation of socially optimally output. With insignificant externalities the absence of a deterrent fee is likely to lead to excessive resource use from society's standpoint. The objective is clear however and the policy instrument is available: various forms of co-insurance can be used to approximate the level of consumption that is characterized by the equality of the value to society of the last unit of service consumed and the real social costs of providing it.

One of the other major objectives of government participation in the funding of health care is the alleviation of the impact of catastrophic health care expenditures. In Diagram IV-4 we illustrate the likely nature of the frequency distribution of annual health care expenditures.¹ The objective of providing protection against catastrophic illness is shown as

1 For an example of a study aimed at describing this distribution is that done by C.H. Berry, Voluntary Medical Insurance and Prepayment, A Study Prepared for the Royal Commission on Health Services, 1964 (Ottawa: Queen's Printer, 1965), and M.S. Feldstein, op. cit., p. 100.

one that involves providing protection against annual expenditures above some arbitrary level, for example \$1000 per year, as shown by the shaded area in Diagram IV-4. Of course there are other alternatives for handling this problem than that of providing complete government funding for expenditures above the predetermined level of the deductible. For example, it would be possible to provide government guaranteed loans repayable at (subsidized?) rates of interest and subject to whatever conditions one thought appropriate to satisfy any distributional goals one felt they should be included as part of this programme.

Diagram IV-4. - Frequency Distribution of Average Annual Expenditures on Health Care



Source: See text.

One of the principle problems with these co-insurance schemes is that the present Canadian environment appears to be one in which the vast bulk of the population might quite likely purchase directly or negotiate as part of their fringe benefit package insurance that would cover health expenditures up to the deductible. Along these lines, the Health Planning Task Force rejected the use of "tiers garant" funding for the present but at the same time urged that further study be made of alternative variants of "tiers garant".¹

Though the purchase of this insurance by consumers might be thought to adequately handle questions of distributional equity, questions of allocational efficiency may well be inadequately treated. Evidence collected by M. Feldstein² for the United States seems to support the expectation that once the consumer is covered by insurance be it privately or government funded the problems of moral hazard are such that the consumer's budget constraint is all but inoperative on his decision to avail himself of health services. In this light, the sharp increases in health expenditures in Canada over the past few years have been nearly matched by similar increases in the United States even though the privately funded insurance in the United States is neither universally purchased nor especially comprehensive in terms of its coverage of all health care.

1 Province of Ontario, Report of the Health Planning Task Force (Toronto: Ministry of Health, 1973), p. 58.

2 M.S. Feldstein, "The Medical Economy", Scientific American (September, 1973), pp. 151-59.

IV. A Research Agenda

Even though one may judge the political climate to be one in which the probability of a change in the funding system in the direction of instituting some variant of "tiers garant" or co-insurance is small, study of the relative potential of these alternatives is warranted. Of the many possible projects that might be mounted in this area and in addition to gaps in our knowledge base already noted, three priorities stand out as especially worthwhile. The first of these involves a thorough study of the extent and nature of direct billing that is currently carried on by a limited number of physicians in Ontario. In addition, this project might involve especially study of the nature of the population thereby served and the nature of the care provided to them in comparison to that provided to the bulk of the Ontario population.

The second line of attack would involve a combined theoretical-applied research project on the implications of the non-price rationing which now characterizes much of the health care sector. In turn these implications could be compared to those expected to follow if one of the variants of co-insurance were introduced.

There is then the question of the interdependence of the system for remunerating physicians and the expected usefulness of the introduction of some variant of "tiers garant" or co-insurance. Most of these last mentioned schemes involve the determination of prices for specific services or treatment packages. The scheme thus requires operational definitions of these and the activities of the physician and his facilities and staff required to render them. Much of the information required to use the fee-for-service

method for remunerating physicians is thereby produced. Further theoretical worked on the expected interdependencies of these systems of paying the physician and charging the patient would thus seem fundamental to decisions to make changes in either of these aspects of health care financing.

Chapter V

The Behavior of Government

With the introduction of hospital insurance and subsequently medical care insurance the importance of the role of government in the health sector increased sharply. There are two aspects of this new role. First, there is the impact that government has as the principal funder of health care and second there is the largely potential role that government has as the single "buyer" of the vast proportion of health care provided in Ontario.

Given this role, the actual and hypothesized behavior of government takes on a significance of equal if not greater value than the other principal economic agents of the health care sector. In this Chapter, therefore we discuss a number of models of the behavior of government and two of the problems to which they might be applied, namely regionalization and federal-provincial cost sharing agreements.

I. General Models of Selected Aspects of Government Behavior

The greater portion of the models developed in Public Finance, the Theory of Governments, and other traditional fields of economists and political scientists are without question relevant to an understanding of government behavior in the health care sector. There are a few of these that seem to be especially relevant and thus are ones we wish to review briefly in order to portray the knowledge base from which we start. Some of the models already discussed in Chapter III with respect to the behavior of the non-profit, non-revenue generating hospital are also very much germane to the present discussion. A closely related body of literature deals with the theory of bureaus and bureaucracies. In addition to this last line of ap-

proach, we also consider in this section bilateral monopoly, the model of countervailing power; the Buchanan hypothesis about government, voter ignorance and learning through experience; the hypothesized fixation of governments with costs without regard to benefits; the theory of optimal grant areas; and the general model of the costs of collective action.

Theory of Government Bureaus and Budget Maximization

Closely related to the models of non-profit, non-revenue generating hospitals discussed in Chapter III and to the literature noted in Chapter II on the factors underlying the formation of "firms", are the theories of government bureaus.¹ From another angle this set of theories is a refinement with a particular goal in mind of the more general theories of bureaucracies especially government bureaucracies which trace their origins back to Max Weber.² Also related are the several works in the general area of organizational theories of the firm.³ A common threat to many of the models of government bureaus is that the bureau, largely through its chief decision maker, is characterized by the pursuit of the objective of maximizing the size of the budget of the bureau.

¹ For example, see K. Acheson, "The Nature of a Bureau", (Carleton University, Department of Economics: Unpublished paper, 1972).

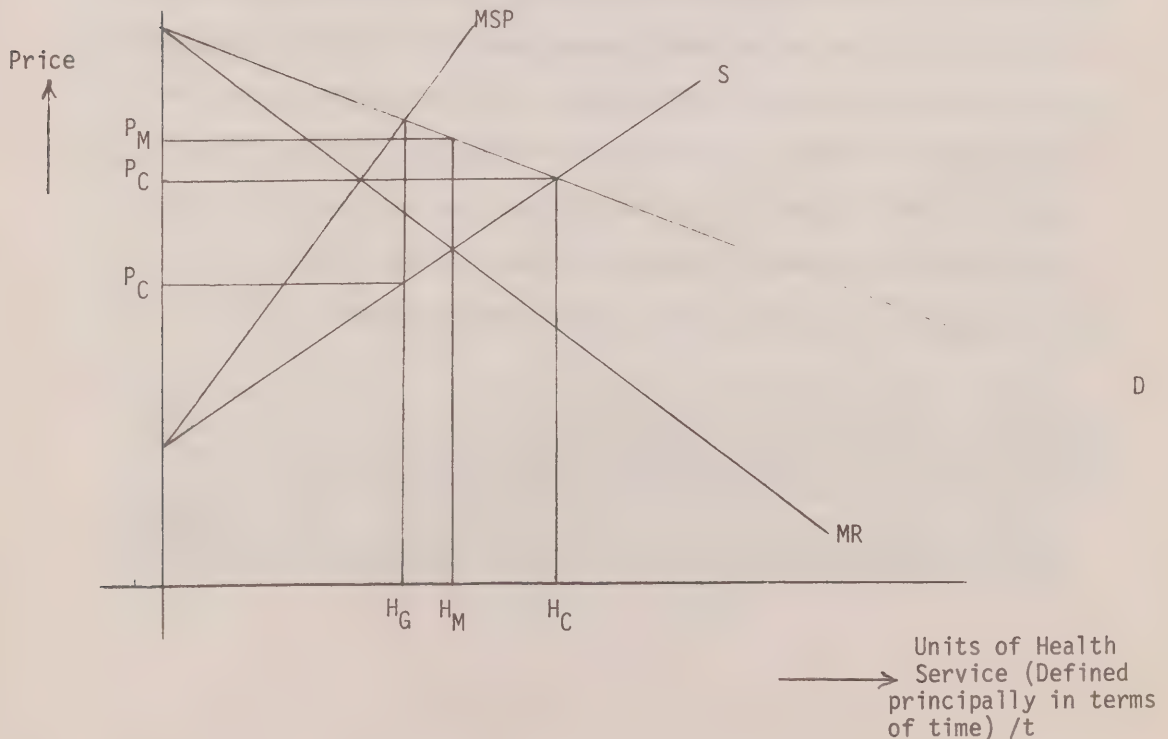
² Max Weber, From Max Weber: Essays in Sociology, translated, edited and with introduction by H.H. Gerth and C. Wright Mills (New York: Oxford University Press, 1946); and Max Weber, Economy and Society: An Outline of Interpretive Sociology, 3 vols., edited, revised and partially translated by G. Roth and C. Wittich with an introduction by G. Roth (New York: Bedminster Press, 1968).

³ For example, see O.E. Williamson, Corporate Control and Business Behavior (Englewood Cliffs: Prentice-Hall, 1970).

Bilateral Monopoly or the Theory of Countervailing Power

Perhaps the candidate from the traditional tool kit of economists most likely to be suggested for analysis of the behavior of government in the health care sector is that of bilateral monopoly. Given that the medical fraternity acting principally through its associations and colleges can be reasonably accurately described as a monopolist seller of health services one might expect that governments entry into the health care field as the principal funder of health care might also lead to behavior on the part of government more consistent with the latent power it possesses as a purchaser of health services on behalf of individual consumers. In Diagram \bar{V} -1 we describe the typical indeterminate solution to a bilateral monopoly model. Physicians are held to want the price-output combination of P_M and H_M while the monoposonist wishes to achieve that of P_G and H_G . Of course, in the case of government, one might well expect the usual goal of the monop-

Diagram \bar{V} -1. - Government as a Monopsonist



sonist to be replaced by the price-output combination of P_C and H_C since it is this combination that would be optimal from society's point of view.

The casual evidence about the behavior of the Ontario government, for example as indicated by the income data discussed in Chapter II, provides little if any support for the contention that government has attempted to move the actual price-output combination from the physician monopoly one of P_M and H_M .¹ The establishment of the Joint Committee on Physician Compensation by the Ontario government would, however, appear to be a step in this direction.

The Hypothesized Fixation of Governments with Costs to the Exclusion of Benefits

We have already discussed the work of T. Marmor and D. Thomas in our discussion in Chapter II of the likely costs involved in a governments attempt to introduce a method of physician remuneration different from the established one providing it has the support of the medical profession. We might recall that the authors had hypothesized that ultimately "doctors will get their way on the methods of their pay". One of the fundamental underlying premises to this hypothesis was the following: "Such governments (those of Western Industrial States), while often disagreeing with physicians and their organizations about desirable methods of payment, prefer gaining medical concessions on the amount of expenditures in exchange for

¹ This raises the question of whether some government departments and agencies might better be viewed as "blocking-type" institutions along the lines developed by A. Breton. See A. Breton, "A Theory of the Economic Council of Canada", in L.H. Officer and L.B. Smith, Canadian Economic Problems and Politics (Toronto: McGraw Hill Co. of Canada Ltd., 1970).

concessions on methods of payment",¹ (the underlining is mine). Similarly, the tremendous concern in Ontario and Canada with the costs of the health care delivery system with seemingly no regard for the benefits obtainable from increased and increasing costs, as for example with proposals to stem the flow of foreign trained physicians, provides further support of the claim that one useful hypothesis about government behavior in the health field is that it may well exhibit a fixation about controlling costs to the exclusion of a concern for the foregone benefits of health services thereby not received.

Buchanan Hypothesis

In an attempt in the mid-sixties to explain the slow but sure squeeze placed by the government of the United Kingdom on expenditures under the National Health Service, J.M. Buchanan had advanced the following hypothesis.² On the assumption that national health care plans, and in particularly the National Health Service, are introduced in part based on the argument that a substantial portion of health care has in fact significant externalities, and further on the premise that the real nature of health care is not closely akin to public goods but rather very much a private good, the population of a given country will slowly become educated to the true nature of health care as they either make use of the health care system themselves or as they witness others taking advantage of it.

¹ T. Marmor and D. Thomas, "The Politics of Paying Physicians: The Determinants of Government Payment Methods in England, Sweden, and the United States," International Journal of Health Services, Vol. 1, Number 1(1971), pp. 73-74.

² J.M. Buchanan, "The Inconsistencies of the National Health Service," Occasional Paper (London: Institute of Economic Affairs, 1965).

In turn, their politicians will begin to feel the pressure of the electorate to control expenditures since the care is really of a very personal kind. Ultimately then governments will begin to place fairly rigid constraints on health care expenditures. Once again, the Buchanan hypothesis seems a reasonable one to entertain especially in view of our discussion in Chapter IV about the nature of health care.

Economic Theory of Government

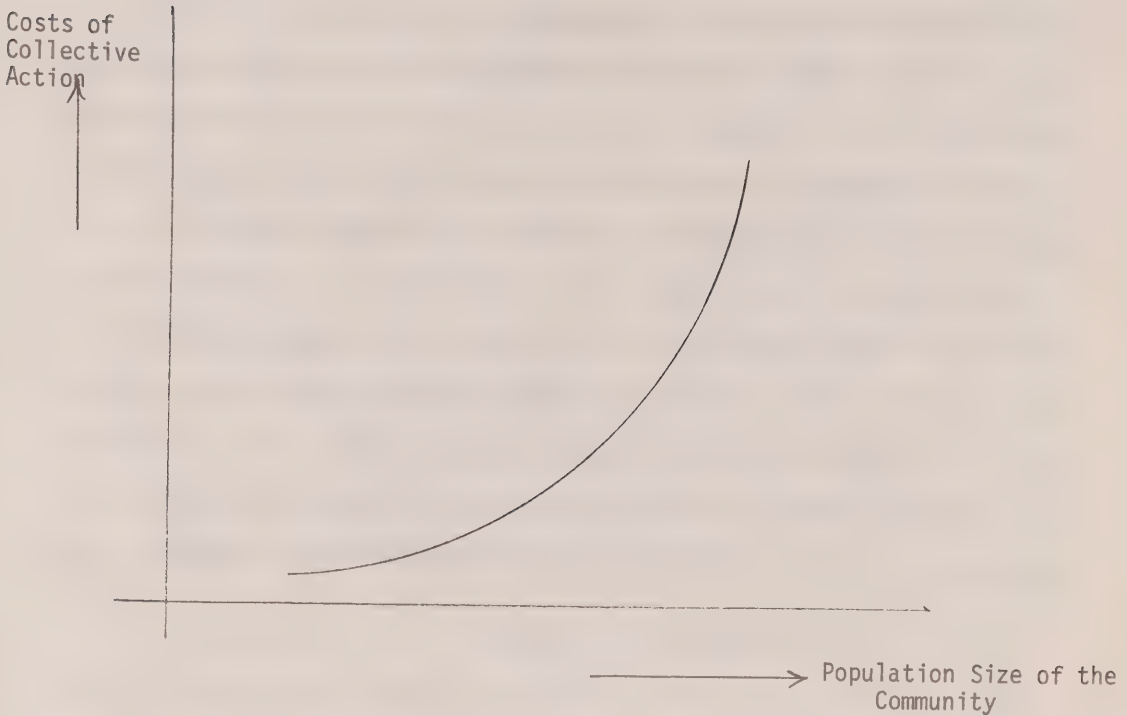
Under the general heading of economic theory of government are the various works that gained their modern day emphasis from the work of J.M. Buchanan and A. Downs.¹ A number of the constructs of more traditional economic analysis, including marginal analysis, have been applied to an analysis of the behavior of governments, especially the more politicized elements of government. An example of this line of attack, we illustrate in Diagram V-2 one of its basic tenets, namely the depiction of the relationship between the costs of collective action and the size of the population being served. In the general case, costs are thought to rise at an exponential rate with increasing population.

Approached from another vantage point, this relationship seems to have been implicit in M. Friedman's general contention² that highly spe-

¹ J.M. Buchanan and G. Tullock, The Calculus of Consent (Ann Arbor: University of Michigan Press, 1962); A. Downs, An Economic Theory of Democracy (New York: Harper and Brox., 1957). For further developments see M. Olson, The Logic of Collective Action (, 1971).

² M. Friedman, Capitalism and Freedom (Chicago: University of Chicago Press, 1962), especially Chapter 9, "Occupational Licensure," pp. 137-60.

Diagram \bar{V} -2. - Costs of Collective Action and Size of Population



Source: See text.

cialized, often geographically concentrated producer groups are likely to find the costs of collective action smaller and the benefits higher than the usually larger number of more geographically dispersed consumers. In the health care area, the net benefits of collective action on the part of physicians appear to have been especially strong. In contrast, those with respect to the collective action of the consuming group as represented by government have been weak if not negative. Within this framework one might well inquire into the socio, economic political characteristics of our society that seem to have yielded this result.

Optimal Grant Areas

One last example of the modeling of the principal elements of resource allocation decisions in a community characterized by a federated, multi-tier government that deserves mention is the work on the relationship between the size of the population covered by the benefits of any particular government funded public good and the nature of the level of government responsible for financing the good either through taxes levied directly or grants received from a higher level of government. A principle work in this area is that of A. Breton.¹ In working out the responsibilities of the different levels of government in the provision of health care from municipal, through regional and provincial to the federal level, the concept of optimal grants as defined principally in terms of the boundary to the spread of benefits is likely to be of not insignificant value.

¹ A. Breton, "A Theory of Government Grants", The Canadian Journal of Economics and Political Science, Vol. 31 (May, 1965), pp. 175-87. See also A.D. Scott, "A Note on Grants in Federal Countries," Economica, Vol. (November, 1950), pp. ; and J. Rothenberg, "Local Decentralization and the Theory of Optimal Government," in J. Margolis, Editor, The Analysis of Public Output (New York: Columbia University Press, 1970), pp.31-64.

II. Regionalization

A. An Historial Sketch of Frequently Observed

The arguments in favor of regionalization of health services within the framework of government funded health care have long been advanced and attempts made to carryout such reorganization. A simple minded thumbnail sketch of the development of the health sectors of many countries includes the following major periods. In the pre-government health plan era, the health care system is indeed regionalized; in fact, within regions it is decentralized (perhaps to the extent of being balkanized) even further with the individual community hospital and the physicians in the community principally serving the inhabitants of the community in question. Coordination of the services within the community may be haphazard while those among different communities are all but non-existent.

In the second stage, characterized the government involvement usually as the principal funder of the care, an attempt is made to integrate across the entire health system each unit of the different broad classes of health care resources. For example, with respect to active treatment hospitals, the role played by each in separate communities is modified by the objective of rationalizing the provision of active treatment hospital care across the entire system. This objective often stems from the historical evolution of government participation itself. It frequently involves participation in the financing of the bulk of hospital care as an initial step, followed by the financing of the bulk of medical care. Subsequently other subsets of care including nursing home care, home care programmes, denticare, pharmicare etc. are added to the range of health care coverage. With each of these distinct steps, the govern-

ment organizes a separate branch to deal with the problems of the particular subset of resources in question. Government likely feels more comfortable with this framework of easily defined sets of responsibilities. The ultimate result is a health care system with distinctly vertical lines of organization.

The third stage in the evolution is characterized by attempts integrate community by community the different subsets of health care resources found in each community. It is at this stage that our knowledge base is in dire need of building. From one view point this third stage of regionalization is an attempt to return to regional level for the determination of priorities. From another, it can be seen as the development of the actual operation of the health care system to a level of integration and coordination significantly more advanced than it was previously. Also involved is the corresponding development of administrative arrangements within the provincial government and between it and the other levels of government including the federal level but more fundamentally the local and, or, regional levels.¹ In what follows we wish to distinguish between the broadly defined functional arrangements of the various elements of the health care delivery system from the administrative arrangements.

B. Alternative Ways of Organizing the Health Care Delivery System

There appear to be several broad lines along with the functional

¹ A description of the Canadian experience in developing regionalized health care systems has been drawn together by Health and Welfare Canada. See "Regionalization of Health Services in Canada - A Survey of Developments," A Working Paper (Ottawa: Long Range Health Planning Branch, Department of National Health and Welfare, February, 1974).

relationships of the elements of the health care delivery system might be organized, two major examples of which are by type of health care resource and by major disease area. The first of these corresponds roughly to "horizontal integration" in the jargon of the industrial organization economist. The present arrangements are an example of this alternative which as we have noted can be explained in part by the historical evolution of government involvement in the health care delivery system. In this arrangement, each individual production unit be it a hospital or the general physicians solo practice are considered along with similar units in communities across the Province. Each such subset of facilities in the provincial system may well be efficiently organized taken as a subset or as a single unit to provide a particular subset of health care goods and services. Coordination, let alone integration, of these subsets within a given community may at the same time be all but non-existent. This is held by the consumer to generate a system of discontinuous care complete with wasted resources through duplication and the absence altogether of certain treatment packages.¹

A second principal method of organizing the health care delivery system and that proposed by the Health Planning Task Force² is more closely akin to "vertical integration". It involves organization along the lines of the major diagnostic illnesses and the total treatment package that is

¹ For example, see, Province of Ontario, Report of the Health Planning Task Force (Toronto: Ministry of Health, 1973), pp. 1-5; and see also E.A. Pickering, Report of the Special Study Regarding the Medical Profession in Ontario (Toronto: Canadian Medical Association, 1973), pp. 81-82.

² Report of the Health Planning Task Force, op. cit., pp. 9-22.

held medically necessary. In this arrangement it is thought to be essential that all the resources directed, for example, to the treatment of cardiovascular problems in a community be involved in a network of working relationships so that consumers with that diagnostic problem can avail themselves of a well integrated, coordinated treatment package. This package in turn may be provided by a variety of health care personnel in a variety of physical facilities. Though a concern for the efficiency of individual production facilities is not ignored, this arrangement emphasizes the efficiency of with which different treatment packages are provided rather than emphasizing the efficiency of the individual production units such as the hospital.

In Figure V-1 we illustrate the alternatives just discussed. The first emphasizes the organization of the health care delivery system along the rows of the matrix whereas the second emphasizes its organization in terms of the columns.

With either of the two principal alternatives, though the principal framework of organization is along one line rather than the other, it is probably essential that the organization along the other line is also considered and necessary networks of communication and, or, operating procedures developed. In other words ultimately a form of "conglomerate" merger may be the most sensible approach for an individual community.

In addition to the two principal alternatives described there may well be others worthy of consideration. What does seem true however is that even with respect to these two alternatives little more than descriptive analysis has been done. Further work exploring the expected impact of

Figure V-1. - Principal Alternative Ways of Organizing the Health Care Delivery System

By Type of Health Care Resource	Cardiovascular Problems	Cancer	Mental Health	Respiratory Problems	Neurological Problems	Gastro-enterological Problems	Ureological Problems	Nephrological Problems	Optthalmological Problems	...
I. Primary Care: 1. Physicians 2. Dentists 3. Community Health Clinics										
II. Secondary Care 1. Active Treatment Hospitals 2. Convalescence Hospitals 3. Chronic Hospitals 4. Rehabilitation Hospitals 5. Nursing Homes										
III. Tertiary Care: Teaching Hospitals										

Source: See Text.

of these systems on the overall health status of the individuals in a community is required.

C. Alternative Ways of Organizing the Provincial Health Care Sector

Though in principle one would wish to organize the administration of the health care system *seriatim* with solving the more fundamental question of the basic nature of the health care delivery system, in practice the main thrust of government appears to have been that of dealing with the administrative arrangements themselves as judged by criteria that are often one remove at least from the health care delivery system.

Of the several ways of organizing the administrative arrangements of the sector, three stand out. These are the present system, one that involves the provincial government as the principal active purchaser of care, and one that involves a significant local and or regional role in defining the priorities of the sector.

The first of these, the current system, is one in which the provincial government has acted in a passive role as the financier of health care services but has not to any great extent "called the tune". Hospitals, physicians in private practice and other health care institutions have largely functioned independently in determining the kinds of health care programmes and treatment packages that will be available in the community, in other words, in answering the basic economic questions of what goods should be produced, how should they be produced and for whom should they be provided.

A second major alternative would be an administrative arrangement

that involved the provincial government in an active role in determining the kinds of health care programmes and treatment packages that would be available in each community. In other words, the provincial government would take on the principal responsibility for directly answering the three basic economic questions of what, how, and for whom. While matching decision making power with fiscal responsibility, this alternative runs the danger of providing a standardized product and one that poorly if at all takes into consideration any differences in needs and preferences for health care programmes and treatment packages as among different communities. The traditional freedom of the physician in diagnosis is not necessarily limited under this alternative. It is the range of treatment packages which he can prescribe that would be constrained.

The alternative advanced by the Health Planning Task Force involves the decentralization of a major part of the decision making to the District Health Council; (it being envisaged that there might well be twenty-five to thirty such councils across the Province). In turn, reporting to each District Health Council might be five or six Area Management Boards. Such boards might take many forms of which two principal ones are the following. First, the Area Management Board might serve as a consolidation of two or more hospital boards. In this circumstance it is argued that it would thereby be possible to consolidate much of the management function of individual hospitals and in turn to effect various economies of scale in the institutional sector of a community's health care system. The second form would see the elaboration of a single hospital's function to include the responsibility for managing the non-hospital ambulatory care sector. The first form of the Area Management Board might well have more potential in

intermediate sized communities with a multiplicity of hospitals whereas the second form might be useful in small communities that have only one hospital or in the more highly urbanized settings that have fairly large hospitals.

There would undoubtedly be many problems involved in working out the precise boundary between decisions that could be made at the District Level and those that would have to remain the responsibility of the Provincial Government. At the same time, the Health Planning Task Force argued that the potential payoff for integrating and coordinating the various components of the health care sector at the community level were so large as to more than offset the problems mentioned above and the uncertainty about how the District Health Councils would work in practice. Not the least of the problems with this alternative is of course the fundamental one of whether decision making can be decentralized if at the same time fiscal responsibility remains with the Provincial Government.

A variant of the alternative that involves decentralized decision making to the local and, or, regional level is one that calls for the transfer to local or regional government the responsibility for the health care delivery system. Financial arrangements would have to be worked out as they are for other sets of services so transferred. The desired result of this variant is that a local government would be in the best position to make decisions on the relative merits of improved public housing, welfare programmes, the health care programmes and so on. In contrast, the frequently heard arguments against this alternative are that the benefits of health care programmes involve a substantial number of crossings by consumers of local government boundaries; that it might be better to

achieve some rationalization within the health sector itself before making decisions about priorities in health care relative to those associated with other goods and services financed and, or, organized by local governments; and that local governments do not at present, and are unlikely to in the near future, possess the administrative talent to take over responsibility for the health sector.

III. Federal-Provincial Relationships

A special problem of intergovernmental relationships exists between the federal and provincial governments in terms of the nature of the constraints placed on provincial governments in order that they take part in programmes for which the Federal government agrees to share in costs. These constraints may be less desirable than otherwise because of the believed need to have agreement by each of the provinces and the federal government on them. In this way individual provinces may adopt what to them is a "third best" health care delivery system. In addition, the establishment of programmes involving cost-sharing often means that a Province is led into these programmes contrary to what it would otherwise perceive to be in the interests of its population because of the significant financial costs of not doing so. From the provincial standpoint both allocational efficiency and distributional equity may be distorted by such cost-sharing arrangements.

There are a number of issues related to these cost-sharing arrangements that are of interest including the matter of the determinants of the choice of any particular cost-sharing arrangement complete with agreed definitions of the resources and, or, goods and services that are considered eligible under the plan. Closely related is the objective of working out the cost sharing arrangement that would satisfy the objectives, stated or otherwise, of the various participants to the programme. There is then from the provincial standpoint an interest in the impact of any chosen cost-sharing agreement on the allocation of health care resources and the distribution of income within the province.

In order to fully sort out the nature and impact of alternative cost sharing agreements, recourse must be had to the models of the behavior of governments and governments agencies discussed earlier in this Chapter as well as to the models of behavior of the other principal economic agents of the health care sector. For example, with respect to the decision as to which health goods and services should be part of federal-provincial cost-sharing agreements, the Breton analysis of optimal grants areas would seem particularly germane. Similarly, in order to understand the agreements made by the provinces and federal government and thus the likely constraints on the range of alternative agreements that fall within the set of admissible alternatives, the economic theory of government would likely prove essential.

Probably the largest body of literature relevant to this general area of cost sharing agreements but especially to the allocational and distributional impacts of alternative schemes is the work of economists on fiscal federalism.¹

In the Canadian context there is relatively little work in the area that is in the public domain.² There is currently work under progress by D. Patriquin³ to analyze the allocational and distributional im-

¹ For example, R.A. Musgrave, "Approaches to a Fiscal Theory of Political Federalism," in Public Finances: Needs, Sources, and Utilization, A Report of the National Bureau of Economic Research (New York: Columbia University, 1961), pp. 97-129; and J.M. Buchanan and R.E. Wagners "An Efficiency Basis for Federal Fiscal Equalization," in J. Margolis, editor, The Analysis of Public Output (New York: Columbia University, Press, 1970), pp. 139-58.

² However, there likely exist studies done by the staff of various government departments.

³ The study constitutes D. Patriquin's Ph.D. thesis for the London School of Economics.

pact on the individual provincial health sectors of the federal-provincial cost sharing arrangements in hospital and subsequently medical care since the introduction of the former during the period 1959 to 1961.

IV. Research Agenda

At least three areas of research appear to have potential significance. These are further theoretical work on the expected behavior of government in its various forms, a longitudinal study of the impact of District Health Councils, and an analysis of the federal-provincial cost sharing agreements.

A major and perhaps primary thrust for research should be further theoretical work in which the models of the behavior of government in its various forms, as discussed in Part I above, were examined with a view to developing a model of the behavior of such proposed bodies as the District Health Councils and the Area Management Boards, for example, as proposed by the Health Planning Task Force. In a similar way one might model the behavior of the provincial government itself in the possible role as an active purchaser-funder of health care. In this research, the methodology of the social scientist, especially the economist in modeling the behavior of economic and, or, political units would seem essential. However, to be useful, such modeling would likely have the highest payoff if carried out on the basis of an adequate understanding of the institutions and institutional framework of the health care sector. A possible thrust of this research would be first to examine two or three of the more frequently proposed models of regionalization, second, to set out several of the crucial policy decisions that would have to be made at the regional level, and third, to hypothesize the behavior of each of the principal economic agents in the health care sector with respect to each decision. In such a way one might be able to simulate the workings of the different models of regionalization

and thereby evaluate the pros and cons of each. Especially important is the issue of whether the monopoly power of producer groups could be effectively countered at the local level by the other actors in the market.

Given that the Province of Ontario has begun to establish District Health Councils in several communities, there would appear to be an especially significant opportunity to evaluate the impact of these Councils on the health care delivery system and especially the health status of individuals in that community. This would involve a fairly comprehensive longitudinal study, the base of which should begin to be provided immediately. This should include a fairly detailed description of the population served, the services rendered, and their health status. Clearly, the data base to be assembled needs careful consideration before launching such a longitudinal study and indeed one might argue that the more theoretical study proposed in the preceeding paragraph should perhaps be completed before this longitudinal impact studies.

Finally, the nature of the federal-provincial cost sharing programme and the impact it has on allocational efficiency and distributional equity both within the Province and across Canada remains an area of potentially fruitful research endeavors of economists.

CHAPTER VI

The Organization of Non-Hospital Ambulatory

Care: Solo vs. Group Practice

I. INTRODUCTION

The one room school house has long since given way to the consolidated district school; the corner grocery store has all but given up to the advance of the supermarket; will or should the solo practice of medicine continue as the principal mode for organizing the non-hospital, ambulatory care component of the health sector? To government, with the principal responsibility for the funding of health care, answers to these two questions are of interest: the "should" because of the potential increased efficiency with which the given bundle of resources in the non-hospital sector might be used and the "will" because of the possibility that the existence of such efficiencies might by itself lead physicians to form appropriate groupings without the intervention of government. Furthermore, since such a small percentage of non-hospital physicians is currently involved in the group practice of medicine, especially in Ontario, should increases in productivity be obtainable, the gains from the move to group practice across the stock of physician manpower would seem to offer potentially large scale benefits.¹

Finally, since there appears to be a growing support for the establishment of groups emanating from various provincial ministries of health and assorted task forces, there is some urgency in proceeding with research aimed at finding answers to these questions. Not only is the process of changing modes of organization from solo to group likely itself to be costly but also once such a shift has been made, the cost of switching back

1. For example, in Ontario in 1971 a net increase in productivity of 2½ percent across the stock of physicians would have been approximately equal to the entire annual output of the province's medical schools in that year.

to a predominantly solo mode of practice would likely be higher still. It thus seems sensible to establish a more solid information base on the potential net benefits of group practice before government intervenes with potentially costly subsidy programmes to help establish the shift to group practice.

When viewed from a distance, the relatively large office expenses of the average fee practice physician, amounting to some twenty thousand dollars in Ontario in 1971,¹ lend strong intuitive support to the hypothesis that economies of scale must exist in the provision of this type of health care. The argument runs along the following lines: "Surely the human capital embodied in the various categories of health care professionals who presently provide or who potentially might provide ambulatory health care could be more efficiently employed in a group setting?" And similarly for a wide variety of now used or potentially usable classes of capital equipment.

This intuitive argument is further bolstered by the observation that we see in other professional occupations the widespread use of non-solo modes of organizing resources as in the practice of law, accountancy, the university sector, and so on.

At the same time, however, we have the apparent fact that the solo practice of medicine continues to be the principal mode of organization, especially in Ontario.

A research strategy thus begins to form. Are there indeed economies

1. Ottawa, Department of National Health and Welfare, Earnings of Physicians in Canada, 1961-71, Health Care Series No. 30 (Ottawa: Department of National Health and Welfare, 1974), Table A4, p. 20.

of scale in the group practice of medicine? Or, if such exist, are they so small as to be outweighed by the average physician's taste for independence? Or, if such exist and are not outweighed by the psychic income associated with solo practice, are the potential returns outweighed by disincentives to individual effort and cost minimization inherent in group practice or by external constraints such as methods of remuneration? The role of the economist in this general area of research must clearly be an important one but especially so with respect to the first question about the existence and extent of any economies of scale; and so in what follows we concentrate on this aspect.

In developing an overall strategy for future research in this area, it is advantageous to set out more precisely the scope of the problem including the distinction between technical efficiency and the concept of general economic efficiency, the definition of the part of the production process to be scrutinized, and the potential sources of economies and diseconomies of scale.

We then proceed to review the general literature in this area and in particular consider the results of a recently completed Canadian study. We then describe two major research thrusts that might be taken in this area, namely that of assembling an appropriate body of data and that of analyzing the performance of existing group practices and clinics.

II. THE PRODUCTION PROCESS

Inputs

The inputs in the form of the health goods and services required to care for a given problem of ill-health may be set out as follows:

1. Inputs provided by the physician contacted initially by the patient:
 - (a) services provided by the physician himself
 - (b) services provided by paramedical and other personnel directly employed by the physician
 - (c) services rendered by the capital plant
 - (d) services rendered by minor equipment
 - (e) services rendered by major equipment
 - (f) services of intermediate products provided or prescribed by the physician.
2. Inputs provided by other physicians within the same medical practice.
3. Inputs provided by physicians in other medical practices.
4. Inputs provided by outpatient departments of hospitals.
5. Inputs provided by inpatient departments of hospitals.
6. Inputs by ^{the} patient himself in terms of the time and effort required to avail himself of the total health care bundle of services:
 - (a) number of health facilities that must be visited,
 - (b) number of visits to each facility.

It is clear that when we set out the production process in this way in an attempt to determine the nature and extent of economies of scale, that we are necessarily involved in considering the whole range of health services that are provided to an individual with a given health problem, both those in the non-hospital setting and those in the hospital setting. This is necessary because the practitioner, solo or otherwise, must be viewed as having at his disposal the hospital facilities in the community as well as the facilities of his own office. Accordingly, it is possible that the mix of services rendered to the patient in the non-hospital ambulatory care sector may well vary closely with the size of the practice in terms of the number

of physicians involved.

It is also clear, I believe, that before we can render judgments about the relative efficiency of the organizational mode itself we will have to be assured that both the quantity and quality of these different inputs into the production process have been accounted for. This seems to be especially important in the case of the physician himself but is probably no less important with respect to the other inputs in the production process.

A particular aspect of quality is worth stressing, namely the impact on quality by the peer review which will by necessity be operative in certain kinds of group practices. Such peer review potentially serves to encourage both the maintenance and use of skills acquired in medical school education and training programmes, and also the acquiring of new skills through continuing education programmes. The depreciation of skills once acquired and the obsolescence of those skills in the face of advancing medical practice and technology would seem to be an omnipresent threat to the quality of care provided by the independently practicing, solo practitioner.¹

Distinguishable from the several components of peer review in the group setting are the potential positive spillovers as a result of peer presence. At least in certain kinds of group practice, the close proximity of colleagues for coffee or lunch time consultation if not indeed corridor consultations would seem to offer significant potential for improvements in the quality of care.

1. Some evidence on the existence of these problems is provided by K. F. Clute in his book, The General Practitioner (Toronto: University of Toronto Press, 1963).

Outputs

In the true tradition of the economist, our consideration of the efficiency of one production process versus another must necessarily involve settling on a useful measure of value added. This is especially true of the health sector in which the range of final outputs or satisfactions produced is extremely large, varying from the alleviation of physical and/or mental discomfort to the saving of life. Now whereas in the case of the economy at large we have market prices to aid in the determination of comparative values of output, for example the value of apple and orange production, in the health sector such prices simply do not exist. Even in the pre-government insurance era it would be difficult to argue that the fees charged for various services accurately reflected relative valuations. This is so because of the strong possibility that the monopoly power of the medical profession with respect to different health goods and services varied so widely, especially in the non-hospital sector.¹ More certain is the inadequacy of fee schedules in use from the onset of government funded insurance since there has been little if any voice of the consumer in the establishment of relative fees either directly or through government and thereby little evaluation of the private or social benefit of the various health services so priced for reimbursement purposes.

There is then the further complication that the value added to any given group of patients through the provision of a given set of health services will vary according to the health status of these patients and their

1. An indication of the wide range of occupational groups that appear to provide competing services to those provided in the health sector is given in K. White, T. F. Williamson and B. C. Greenberg, "The Ecology of Medical Care," New England Journal of Medicine, Vol. 265 (1961), p. 885.

perception as to what constitutes adequate or good health care. This will be so in spite of the fact that the fee paid for the services does not itself change. Thus, for example, the solo practitioner in one community whose citizens might have relatively low standards of what constitutes good health care may well get by with an office visit that involves no more than a slap on the back, a "How's the family?", and an "Everything's just fine!" (or, in other words, Mother Nature will work her magic!). Whereas in a community with higher standards, adequate care as perceived by the patient might well involve a team of physicians in a group carrying out a rather extensive battery of diagnostic tests and treatment packages. To the extent that group practices that involve taking advantage of economies of scale will likely be more often found in the larger communities, there may well be an implicit bias in favor of the solo practice on these grounds.

From what has been said in the immediately preceding paragraphs about the quantity and quality of inputs and of problems of measuring output, it should be clear that simple measures of output will not capture the sense of "value added" that is necessary in order to determine the relative efficiency of various modes of practice. Such measures as the number of patient visits, the volume of various services provided, the revenues of the physician, etc., simply are not satisfactory.

III. GENERAL MODEL OF THE RELATIONSHIPS BETWEEN REAL RESOURCE COSTS AND THE OUTPUT OF HEALTH CARE

The Model

Once having agreed on the nature of the production process and thereby the principal elements of it that must be considered, we may turn

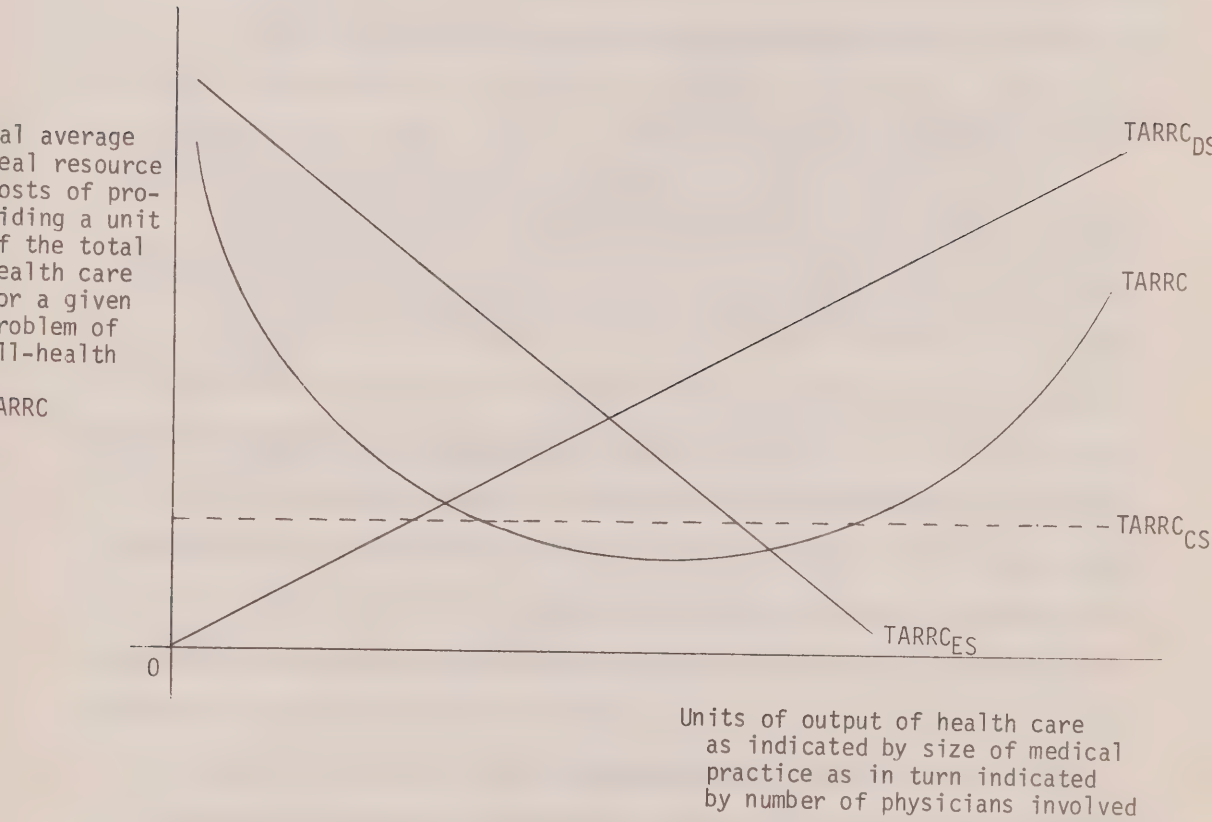
to the nature of the relationship between real resource costs and output. In the diagram below we illustrate the three principal types of relationships which we might expect to obtain.

The principal alternative hypotheses about the relationship between total average real resource costs and size of medical practice are as follows:

(a) Hypothesis H_U , that the average cost-size relationship exhibits a traditional U-shape as indicated by the curve $TARRC$, that is, that there are economies of scale in the first range followed successively first by a range of constant returns to scale and then by diseconomies of scale as size of practice grows still further; (b) Hypothesis H_{ES} , that the average cost-size relationship exhibits economies of scale throughout the observed range of size of practice as indicated by the curve $TARRC_{ES}$ below; (c) Hypothesis H_{CS} , that average costs do not vary as size of practice increases as shown by the line $TARRC_{CS}$ below; (d) Hypothesis H_{DS} , that the average cost-size relationship exhibits diseconomies of scale throughout the observed range of size of practice as indicated by the curve $TARRC_{DS}$ below.

At this point it is well to stress that the framework laid out above is concerned solely with the matter of technical efficiency; that is, the relationship between real resource costs per unit of output and size of productive facility. It must be emphasized that the question of economic efficiency is not comprehensively addressed by examining the nature and extent of economies of scale. Even if, for example, continually decreasing costs are found to exist as size of practice increases, it does not follow that solo practice should be abandoned. There may well be a significant proportion of communities in which the total demand for health care that is to be met simply does not justify the presence of more than a single physician once

DIAGRAM VI - 1 - Alternative Production Functions



Source: See text.

one accounts for the net social benefits of services to the patient which in turn must involve an assessment of transportation costs (in time and money) on the patient's health and general well-being. This last factor should not be overly stressed in view of the consumer's revealed preference in other sectors of economic activity, if not in the health sector itself, for taking on fairly large transportation costs in order to avail himself of the shopping facilities of particular (usually larger) communities.

Potential Sources of Economies of Scale

The potential sources of economies of scale in the provision of non-hospital ambulatory services derive readily from an evaluation of several inputs of the production process. These sources may be set out as follows: (a) economies associated with the specialization and division of labour among non-medically trained personnel such as those performing secretarial, clerical and managerial functions and equally, if not more important, among various kinds of health care personnel including the physician himself because specialization has been so highly developed; (b) economies associated with the capital plant and in particular with minor and major equipment often found in solo and group practices; (c) economies associated with the provision of higher quality services rendered by personnel of given abilities because of interaction and consultation among physicians and paramedical personnel in group as opposed to solo practice; (d) economies associated with the provision of a larger fraction of the total health care package required to meet a given health problem being provided ^{to} a patient in a group practice setting as a result of ^{the} physician of initial contact being able to consult immediately with his colleagues; there is thus the likely saving of

the patient's time because of fewer visits being required for each health problem (that is, more services are rendered per visit); (e) economies associated with the more efficient, less costly flow of information between physician of initial contact and specialist colleagues in the same practice for all those cases in which referral takes place among physicians of the given group practice.

There may well be other sources of economies of scale of a more dynamic nature related to the use of new innovative techniques. The major possibilities would seem to be the following: (a) use of new managerial administrative techniques are more readily adoptable in group practice setting because of ^{the} possibility of hiring an administrative manager; (b) use of new types of paramedical personnel also ^{is} much more easily developed in a group practice setting.

Potential Sources of Diseconomies of Scale

The principal source of diseconomies of scale would seem to be the traditional one of problems related to the management of personnel. This may be particularly so in the medical field in which the education and training of the senior health professional, the physician, has been predominantly influenced by the solo practice mode.

There is then the further source, which is related to the first, that involves the efficacy of methods of enforcing work-effort among physicians themselves and similarly of enforcing cost-minimizing resource use patterns of activity among each of the physicians of the group. This problem of slippage or of X-inefficiency will clearly be related to the methods whereby any group of physicians decide to divide the cash flow to their entire group.

One source of productive effort lost to the health sector that should not, I believe, be considered a diseconomy of scale is the decision a physician makes on allocating his time between leisure and income even when this decision appears to be closely related to the type of practice mode. If the group practice mode does permit each physician to better organize his work time and in particular to follow a work schedule more characteristic of the average worker, the reduction in his total productivity with respect to the output of health care must be counter balanced by what presumably are increases in the society's sum of satisfactions because of his increased leisure. This consideration seems to suggest that our measure of input in the productivity analysis should be per hour unless we are prepared to evaluate increases in leisure time.

Additional Problems of Estimation

Given the foregoing discussion, the information required about real resource costs as opposed to money costs will also be hampered by the following: (a) Government or other reimbursement schemes may reimburse on the basis of a fee-for-service schedule that does not differentiate between the costs of a service provided by the physician and the costs of the same service provided by paramedical personnel in the employ of the physician. Group practices may therefore be characterized by a lower real social cost of health care delivery which will not be reflected in lower money costs. (b) The quality of the total health care provided by medical practices may well vary with size of practice; such variations, which themselves are likely to be reflected in such figures as "average income (gross or net) per physician" and "average cost per patient," to account for the bigger bundle of services

provided, will all but preclude use of such expenditure items as part of the evaluation of the alternative hypotheses set out above. (c) The ability of medical practices to obtain monopoly and/or monopsonistic profits may well vary with size of medical practice causing even further distortions between real resource costs and money costs.

IV. LITERATURE SURVEY

A survey of the North American literature in this area has recently been completed by M. Prime and myself.¹ In addition to reviewing the conclusions drawn as a result of this survey, work of several principal researchers in the United States and Canada, including R. Bailey, U. Reinhardt, J. Newhouse and R. G. Evans et al., will be briefly examined.

R. M. Bailey

There are three related studies by Bailey that we might consider. In the first of these² the productive activities of some 70 physicians in various sizes of groups practicing internal medicine in the San Francisco area are examined. The average revenue and profit data analyzed seemed to be consistent with the hypothesis that economies of scale were present. However, a more detailed examination of the sources of revenue, undertaken in order to determine the similarity of the output mix among the practices, revealed substantial differences. Bailey's attempt to account for these differences involves the use of proxy measures of output such as the actual number of units of various categories of services produced. Though inadequate

1. R. D. Fraser and M. Prime, "Economies of Group Medical Practice: A Survey," Queen's Discussion Paper Series, No. , forthcoming. In this section I have drawn heavily on this work.

2. Richard M. Bailey, "Economies of Scale in Outpatient Medical Practice," Group Practice, Vol. XVII (July, 1968), pp. 24-33.

for the purpose, this further analysis led Bailey to conclude that solo practitioners had higher levels of productivity than physicians in groups.

The general conclusion reached by Bailey in his second paper¹ is again that internists in group practice were more inefficient than their solo-practicing counterparts. In this study he reported that solo and group practice internists seemed to use paramedical personnel to roughly the same extent. That the output of group practice internists in terms of the volume of different broadly defined categories of services was lower than that for solo practitioners was explained in terms of the nature of the leisure-income trade-off made by individuals in these two groups.

In a similar vein Bailey concludes in his third paper² that solo practitioners are as productive as those in group practices. His concern in this paper, having emphasized the distinction between the services of the physician and those of his ancillary personnel, is with the number of patient visits per hour and total patient visits per period.

Bailey's concentration on fee-for-service practice and on internists, though limiting the scope of the conclusions that can be drawn for the whole of the practice of non-hospital ambulatory care, does have the advantage of standardizing his analysis for two potentially significant factors. At the same time, however, he has not been able to adequately consider the quantity and quality of the range of inputs. Similarly the use of revenue

1. Richard M. Bailey, "A Comparison of Internists in Solo and Fee-for-Service Group Practice in the San Francisco Bay Area," Bulletin of the New York Academy of Medicine, Vol. XLIV (November, 1968), pp. 1293-1303.

2. Richard M. Bailey, "Economies of Scale in Medical Practice," in Empirical Studies in Health Economics, edited by H. E. Klarman (Baltimore: The Johns Hopkins Press, 1970), pp. 255-273.

or income data as a measure of output is simply not good enough. As well, the sample size was so small as to severely limit the conclusions that could be drawn.

U. E. Reinhardt

There are three principal papers prepared by Reinhardt that we might review briefly. In the first of these¹ regression techniques are used to estimate the possibilities for substituting the services of paramedical personnel for those of the physician as the size of group practice increases. The principal data base used by Reinhardt is mail survey data of physicians in the United States as obtained by Medical Economics Inc. Controlling for factor inputs, he finds that group practice is characterized by increased productivity.

Reinhardt's second paper² again involves the estimation of production functions for U.S. physicians. Again his data base is that stemming from the Medical Economics mail survey. And once again his conclusion is that single specialty group practice is more efficient than solo practice. In particular, he concludes that the average American physician could increase his productivity some 25 percent by employing roughly twice the number of paramedical personnel.

Finally, Reinhardt has produced a joint paper with D. E. Yett in

1. U. E. Reinhardt, "Physician Productivity, the Supply of Physicians Services and the 'Physician Shortage in Canada,'" a paper presented to the Fifth Annual Meeting of the Canadian Economics Association, St. John's, Nfld., June, 1971.

2. U. E. Reinhardt, "A Production Function for Physicians Services," Review of Economics and Statistics, Vol. LIV (February, 1972), pp. 55-66.

which the authors report on their estimation of production functions using the same source of data as in Reinhardt's previous papers.¹ Their prime purpose is an attempt to disentangle the effects of factor inputs from type of practice mode. The conclusion they reached with respect to the latter variable varied from Reinhardt's earlier work, namely, economies of scale in the group practice of medicine were not generally found to exist.

Again, as in the case of Bailey's work the research work accomplished by Reinhardt advances the field somewhat. In particular, a much more successful attempt is made to account for the range of inputs, their quantity and quality. The same cannot however be said of the output side.

J. P. Newhouse

In a principal paper² in this area of research Newhouse advances two hypotheses to explain the apparent lack of observed economies of scale in the practice of group medicine and then analyzes a body of data to check the consistency of the data with the advanced hypotheses. The first of these is that cost and/or revenue sharing as occurs in groups is characterized by problems of X-inefficiency. Incentives to use least cost methods of production are much less than in solo practice unless some method of enforcement can be introduced and in any case such enforcement methods as are introduced are not in themselves costless. Similarly, under revenue sharing there is likely to emerge a problem of shirking on the part of the physicians

1. U. E. Reinhardt and D. E. Yett, "Physician Production Functions Under Varying Practice Arrangements," Technical Paper No. 11, Community Profile Data Center (Washington: H.E.W., August 1972).

2. Joseph P. Newhouse, "The Economics of Group Practice," The Journal of Human Resources, Vol. VIII (Winter, 1973), pp. 37-56.

themselves. His second hypothesis then relates to the decision each physician faces with respect to allocating his time between leisure and income; in particular, in the group practice setting the physician may well find it easier to spend more of his time in leisure activities than in earning income. Thus, his overall expectation is not that economies of scale do not exist but rather that these effects will be more than outweighed by these other two factors.

Newhouse examines in detail a rather small sample of eleven solo practitioners, nine assorted group practices and three clinics. He finds significant differences in overhead costs which are consistent with his X-inefficiency hypothesis. The same is found for salary costs. With respect to his hypothesis about the leisure-income decision, his sample data did not lend support.

R. G. Evans et al.

In a recently completed study¹ R. G. Evans, E. M. Parish and F. Sully have analyzed gross billings data for the British Columbia Medical Services Commission for 1969-70. Regression techniques are used to estimate the excess in earnings, by specialty, of group practice physicians versus their solo counterparts. The effects of location, size of group, physician density and age are controlled for in the regression. The conclusions drawn are that with significant differences among specialties, group practice is characterized by modest gains in productivity. It should be noted however that the various inputs to the production process other than the physician's services

1. R. G. Evans, E. M. A. Parish and F. Sully, "Medical Productivity, Scale Effects, and Demand Generation," Canadian Journal of Economics, Vol. VI (August, 1973), pp. 376-93.

are not accounted for. At the same time with respect to the physician, both his specialty and age are considered. The output side of the equation is subject to problems we have already discussed, many of which are noted by Evans et al. themselves.

R.D. Fraser and M. Prime

A review of the literature¹ in the area of the nature and extent of economies of scale in the group practice of medicine leads to two conclusions. First, the weight of the evidence gathered thus far suggests that economies of scale in the true technical sense exist at least on a modest scale. The word "evidence" is underlined to distinguish it from the several and often advanced a priori arguments. An evaluation of the latter leaves one in a quandary since on the one hand the traditional arguments that suggest the existence of economies of scale, are balanced by those relating to the possibilities of X-inefficiency. Similarly, we may qualify the phrase "at least" by stating that the measures of output commonly used probably bias the analysis against the finding of economies of scale.

The second conclusion is the "weight" of the evidence is not as weighty as it probably could and should be for the guiding of public policy. The area is still comparatively unresearched even though it does present potentially significant gains to society.

It should be emphasized that the bulk of the physician groups studied in the literature just reviewed could best be described as loosely vertically or horizontally integrated groups of independently practicing physicians. To this extent, then, the results do not shed much light on the magnitude of the benefits that would be forthcoming were physicians to be organized in groups whose prime

1. R.D. Fraser and M. Prime, op. cit.

goal was to maximize the amount of health care provided by the group and thus in groups in which the team concept of health care delivery was pre-eminent and especially in which the reimbursement mechanism was geared to facilitate pursuit of this prime goal.

V. A RESEARCH AGENDA

Two major thrusts for research in the economics of the group practice of medicine seem especially worthwhile developing. The first of these would direct researchers to a micro type study of the actual production processes found in group practices. The second thrust would be directed towards the establishment of an adequate data base. Though one could argue for some time on the pros and cons of which of these thrusts takes precedence over the other, I propose to argue that they go forward at the same time. The micro firm type studies of on-going practices are necessary to familiarize the researcher with the actual production processes so that the several inputs and the components of output can be identified and decisions made as to which data are required in order to carry out the appropriate research. At the same time there is a base of experience, largely the result of recently completed U.S. studies using the data generated by Medical Economics through its mail survey of physicians, that is likely sufficient to permit this type of data formation to begin at once.

Data Assembling

In order to carry on the kind of studies that are required to answer the basic policy issues facing government with respect to the encouragement of group practice there is the necessity of developing an adequate data base on the inputs and outputs of the production processes in the provision of non-hospital ambulatory care. Towards this end, there would seem to be two potential directions one might go. First, it is conceptually possible for governments to develop a health information system based on their hospital and medical care in-

insurance records that would permit fairly thorough analyses of issues related to economies of the group practice of medicine. Secondly, one could follow the lead of Medical Economics, Inc., and develop a data bank based on the answers to mailed surveys.

In all of Canada only two or three provinces have set up their insurance records in a manner that permits the linking of hospital use by patient and physician to the use of non-hospital ambulatory services. In Ontario at the present time this type of health information system does not exist. Furthermore, it does not appear to be imminent.

The mail survey has all the problems associated with interviewing as a source of adequate data. The answer to "How many hours do you work per week on average?" and "What was your total gross income last year?" are likely characterized by only a rough degree of accuracy. However, as long as questions require factual answers as opposed to statements about motivations, the data base so collected would likely be satisfactory for the intended purposes.

In order to establish an effective mail survey, the cooperation if not the participation of the C.M.A. and/or the O.M.A. would be required. Because of the lead of their U.S. counterparts, this might not be such an insurmountable obstacle. This participation would seem to offer especially valuable help not only in securing the cooperation of the surveyed physicians but also perhaps in obtaining some expertise on information that might be collected that would assist in the evaluation of the output side of the equation.

In developing the questionnaire, the model provided by Medical Economics, Inc., should of course be given serious consideration both in

order to take advantage of the effort that has already gone into its development and also to facilitate inter-country comparisons. Equally important would be consideration of the questionnaire used by S. Judek in his 1964 Royal Commission Study¹ and that used by the C.M.A. in their 1967 survey of Canadian physicians.²

Once gathered, these data should be useful for two broadly different lines of research. First and perhaps foremost, a description of the extent and nature of group practice could be drawn together to fill a present gap in our data base. It could then hopefully be examined together with the Judek and C.M.A. data to determine if any historical trends appear in the last ten years. This type of survivor analysis³ could, if well done, offer some insight into the assessment of physicians themselves of the net advantages of group practice as revealed in their decisions to form groups of particular kinds.

Of course the second line of research that would likely follow immediately on the availability of such a data base being established is the estimation of cost and production functions using econometric methods. In this direction not only should new avenues of analysis be explored but also some of the studies recently completed in the United States might

1. S. Judek, Medical Manpower in Canada, A Study Commissioned by the Hall Royal Commission on Health Services, 1964 (Queen's Printer: Ottawa, 1964).

2. Canadian Medical Association, Group Practice in Canada (Toronto: Ryerson Press, 1967).

3. One such study has recently been completed for the United States. See Paul B. Ginsburg and H. E. Frech, "Optimal Scale in Group Medical Practice: A Survivor Analysis," Workshop Paper No. 7207, Michigan State University, Econometrics Workshop, January, 1973.

with advantage be reworked using Canadian data.

Micro Firm Studies

There are presently operating in Ontario several group practices and clinics which to this point in time have not been thoroughly studied. Among these are the Sault Ste. Marie and St. Catherines Clinics, the St. Mary's group practice and several others. Data describing the two group practice clinics have been gathered from time to time but have not yet been subjected to a thorough-going analysis of the kind required and as outlined in the earlier parts of this paper. Because of the problems of linking hospital use data to non-hospital data, these micro studies will necessarily have to involve the researchers in a fairly extensive task of assembling data.

In addition to the utilization data, the data now coming forward from the 1971 Census of the Population, especially that compiled for Census tracks, could be with advantage used to control for the differing social, economic and demographic characteristics of the populations served by the practices in question. Similarly, data on vital statistics for individual communities should be used to begin the process of assessing the impact of various forms of practice on health status.

Since the survey data discussed above is not likely to yield the information required to adequately handle the output side of the equation, one would, I believe, look to these micro firm type studies to begin progress in this direction.

CHAPTER VII

The Substitutability of Human Health Care Resources*

I Introduction

The possibilities for manpower substitution (principally between broadly qualified and more specifically qualified health care personnel) have been demonstrated in many countries throughout the world partially through recent trends of manpower utilization. That these trends and traditions are not present in Canada's health sector has likely magnified the so called "manpower crisis" that we face today.¹ Thus, by way of introduction to this section, we will identify some of the approaches to manpower substitution from an international point of view.

The Russian Academy of Science imported the feldsher (field barber) from Germany in 1700. Since this time the feldshers, working as physicians' assistants in urban areas and as the primary source of care in rural areas, have been firmly established in the delivery of health services in the U.S.S.R.² Despite expansion of physician training programs in Russia in the 1930's, the persistence of personnel shortages especially in rural areas motivated expansion of feldsher training. In 1965 there were about two feldshers for every physician in addition to many other types of auxiliary personnel.

1. "Manpower crisis" is a phrase more popularly used in U.S. than in Canada. I suspect that this is only true because the government and the unwary taxpayer have been more willing to pay the price for wasteful manpower utilization here than in the United States.

2. V.W. Sidel, "The Feldsher in the U.S.S.R.", Annals of the New York Academy of Science (December 31, 1969), p. 987.

* The research assistance of Mr. T. Edwards, an Honours B.A.(Economics) Graduate of Queen's University, in assembling the background material for this Chapter and in setting out the first draft of it, is gratefully acknowledged.

A more familiar example of substitution seen in a comparison of international patterns of health care delivery is that of the nurse-midwife. Hellman notes that the majority of the world's babies are still delivered by midwives and that the North American obstetrician has, in fact, been a very poor substitute for the comprehensive and well-distributed care that midwives provide in both highly developed and underdeveloped countries.¹ Other culturally sanctioned health personnel that have filled in for unavailable professional care are the "barefoot doctors" of China² who provide primary care on the basis of only six months of training and the native healers of India who are widely accepted on religious grounds.³

In the last decade we have witnessed the growth of a new group of health care personnel who must be recognized for their potential in helping to meet the planned objectives of the health sectors of many countries. For the most part this group has been made up not of new types of personnel but rather have been auxiliary workers in the more traditional occupations. The work however now accounts for an ever increasing proportion of the total that provided in the health sector. The individuals in question are the growing numbers of nurses, technicians, and unskilled labourers; indeed, this

-
1. Louis Hellman, "Nurse Midwifery", Annals of the New York Academy of Science, (December 31, 1969) p. 896.
 2. V. W. Sidel, "Discussion of Delivery of Care: Acute and Chronic Illness", in The Greater Medical Profession, The Josiah Macy Jr. Foundation, Rockefeller Plaza, New York, 1973) p. 72.
 3. Fredrick C. Redlich, "Delivery of Care: Mental Disorders", Dr. Redlich notes that in India there are 300 psychiatrists and well over 100,000 native healers, to point out the alignment of most available and most effective methods when intervention is not a problem. His article is published in The Greater Medical Profession (The Josiah Macy Foundation Rockefeller Plaza, New York: 1973) p. 98.

growth has been sufficient to radically change the ratio of professionals to non-professionals. In the United States, a rough estimate of the changing ratio is that in 1900 professionals and sub-professionals were of nearly equal numbers whereas in 1966 there were between 5 and 8 non-professionals for each professional.¹ The change has come quite recently according to Weiss who estimates that for 1950 to 1960, 117 thousand low-skilled jobs were substituted for 113 thousand high- and middle-skilled jobs.² Similarly in Russia the growth of allied health personnel in the last forty years has caused the proportions of professionals to change from 1 to 1.5, to 1 to 4.³ Since the initiation of the British national health system 28 years ago, the number of general practitioners has increased by 25 percent while the number of nurses has doubled and other technical staff has increased by 150 percent⁴. Finally in Sweden a growth of health personnel from 1950 to 1965 has been 65 percent with the most significant gains coming in the non-professional categories.⁵

As for Canada, perhaps the most important traditions have been those that work contrary to efficient manpower utilization. These are the

-
1. H. I. Greenfield & C. A. Brown, Allied Health Manpower Trends & Prospects (New York: Columbia University Press, 1969) p. 31.
 2. R. Fein, The Doctor Shortage: An Economic Diagnosis (Washington: Brookings Studies in Social Economics, 1967) p. 118.
 3. V. W. Sidel, "The Feldsher in the U.S.S.R...."
 4. Sir George Godber, "Manpower Problems on Either Side of the Atlantic", The Greater Medical Profession (Josiah Macy Jr. Foundation, Rockefeller Plaza, New York, 1973) p. 23.
 5. Gunnar Wennstrom, "Training of Health Workers in The Swedish Medical Care System", Annals of the New York Academy of Science (December 31, 1969) p. 985.

substitution of more-expensive for less-expensive manpower; e.g., obstetricians for midwives, the continuing support of an expensive licensure system and the use of a reimbursement system that offers few economic incentives for efficient operation. Accordingly the growth of auxiliary personnel has not been sufficient to reveal a pattern of substitution comparable to that of other industrialized countries. For instance, the ratio of dental hygienists to licenced dentists in Ontario was 1 : 9 in 1967 and 1 : 7 in 1972. Interestingly, the growth rate of dentists exceeds that of hygienists, even though each dentist could provide work for at least one hygienist.¹ So, too, for physicians, the rate of growth in their numbers over the last five years in Ontario has averaged 7.9 per cent. This is greater than the growth rate of nurses and orderlies and almost as large as that of nursing assistants. Not only have the usual types of manpower not been substituted but neither have new types. Beyond a few demonstration projects, physicians assistants and pediatric nurses are almost unknown. In addition, it should be noted that the distribution of health personnel in Canada is by no means uniform with respect to population. From a regional point of view, the range of dentists per capita in 1972 was from 1 : 8, 313 in Newfoundland to 1 : 1,076 in Prince Edward Island to 1 : 586 in Ontario.² These rough figures suggest that Ontario and Canada residents are paying excessively for health care by supporting inefficient production methods and that the care received could be better distributed.

1. Profiles on Health Occupations, Ontario Department of Health Research and Planning Branch, 1969 and Canadian Manpower Inventory 1973, Dept. of Health & Welfare, Canada.

2. Canada Manpower Inventory 1973, Dept. of Health & Welfare, Canada.

In the following section we will be chiefly concerned with the possibilities for improving the efficiency of production methods. The organization of this section will be to first identify the production function for health care, and to discuss the nature of health care input and output. Secondly, we wish to discuss potential areas for productivity growth and the constraints to such growth. Thirdly, we look into some of the major policy issues to which economic analysis can be readily applied. Fourthly, some selected case studies of substitution possibilities will be reviewed with the intention of demonstrating the potential for substitution. Finally some suggestions regarding research proposals will be made.

II Production Function Approach

In this section we review the possibilities for manpower substitution in the production of health services. From an economist's point of view this involves an analysis of the health sector production function, focusing on manpower productivity. A generalized production function for the health sector might be written:

$$Q = f(L_1 \dots L_n, K_j \dots K_n, T)$$

where Q is the quantity of health services produced by various combinations of L_i , health manpower inputs (physicians, nurses, technicians), and K_j , capital inputs (buildings, equipment, disposables) at a given level T , of

technology.¹ The productivity of any input; i.e., the output-input ratio, can be changed:

- a) by increasing the proportion in which the more-productive factors are combined with less-productive ones,
- b) by changing the technological level embodied in the production function,
- c) by taking advantage of economies of scale.

Our current interest is with the first case of potential productivity gains through substitution of factors.

The general issues involved in the analysis of input substitution are whether or not technical possibilities for substitution exist and if they do, are they economically advantageous. Technical feasibility relates to the potential for producing particular physical units of output using alternative resources or combinations thereof. Economic feasibility relates to question of whether a proposed activity, that has been already demonstrated as being technically feasible, is one for which the value of the output exceeds the cost of carrying out the activity.

Prior to an evaluation of suggested substitutions, we should consider the general theoretical concepts of health sector inputs and outputs and the methods used for accurately measuring these. Manpower inputs are usually clearly defined by the legal and administrative bodies that supervise their activities. By assuming constant productivity and quality over each category of manpower it is possible by head count to develop a fairly accurate measure of labour input. So, too, the capital input is observable in

1. For a complete discussion of production function analysis of the health sector see U.E. Reinhardt "Manpower Substitution Research", Health Services Research Journal (Fall, 1973) pp. 200-227.

standardized units although there remains the difficulty of differentiating technology changes and capital substitution. By weighting the physical inputs by their price, it is possible to calculate the economic productivity of each input.

The definition and measurement of output, however, are far more ambiguous. Traditional measures such as number of patient visits, procedures completed, cases treated, prevalence reduced and similar approaches are all inadequate, though widely used, in so far as they measure the process rather than the output of health care. These intermediate variables, however, are the ones most readily quantifiable given the availability of data. Also, if output is measured by an index of health status improvement, it may be impossible to distinguish between variations in environment, patient contribution, and input productivity. If the data and resources exist, by all means, real health improvement is the preferable output yardstick.¹ But in less ambitious studies intervening process variables are used. These can be made more useful by measuring a variety of dimensions: content or service intensity, quality and purpose. Purpose is often dealt with on broad criteria of specialty and job function, while service intensity is used for detailed examination. Kovner developed the "Identifiable Medical Procedures" approach to analyze service intensity that breaks down the activity of labour units into specific

1. For an excellent discussion of output measurement see G. W. Forrance, "A Generalized Cost Effectiveness Model for The Evaluation of Health Programs", Research Series 101 Faculty of Business, McMaster University. By same author mimeograph, "The Health Status Scale: An Output Measure for Productivity".

comparable functions.¹ These can be weighted by some scale, (for example, by the California Relative Value Scale) and accumulated to produce a detailed picture of the real manpower input of each visit according to the number and importance of identifiable medical procedures. This method can be expanded by evaluating each identifiable medical procedure with respect to the level of training and liability involved such that each procedure can be assigned to the least expensive manpower input.

Thus far nothing has been said of the quality of input or output. Many physicians, hospital administrators and other health professionals feel that attaining the highest quality of care for all should be one of the objectives of the health sector and public policy. This is understandable given the fact that the health sector can be described as an "anxiety driven" system. Most people seek medical care because of pain, possible disability or death or emotional discomfort. At the same time, health professionals are imbued with the moral, ethical, legal and human consequences of their actions. Further, some recent and spectacular breakthroughs in medicine (heart transplants and hemodialysis, for example) have greatly increased public expectations of both the efficacy and availability of medical care.

Unfortunately, many do not realize that medical advances have been less than dramatic in other fields, that there is still much about disease etiology, growth, and transmission which is not known and that many of the medical miracles are extremely expensive and not always clearly beneficial.

1. J. W. Kouner, "Measurement of Outpatient Visit Services", Health Services Research Vol. 4, No. 2 (Summer, 1969) p. 113.

If one takes the "highest quality of care" to mean the best techniques and highest professional standards currently operative (Rolls Royce Medicine) then the cost implications of extending this level of care to the entire population are virtually infinite. With respect to the technology level utilized it is becoming increasingly clear that with sufficient research funding many scientific obstacles can be overcome. Thus, there is no limit to the quality of health care (broadly interpreted) that could be provided unless the full implication of a scarcity of health resources is understood.

Determination of a minimally acceptable level of quality will depend not only on professional standards, but also on the value members of our society assign to health and the quantity of resources they are willing to allocate to health services. It also must be remembered that in the world of scarce resources there exist trade-offs between the distribution of medical care, the total quantity of services and the quality of care. From this milieu of social-political decisions the point relevant to productivity analysis is that an increase in distribution or productivity may well come at the price of reduced quality. The greatest quality for the greatest number contains one greatest too many.

In addition to the value judgements associated with decisions regarding the quality of care, there are also the technical problems of how to incorporate quality levels and quality changes into an analysis of productivity. Capital input is fairly well defined in terms of its technical

quality but its benefits in terms of its impact on the health status of the patient is as difficult to assess as is the impact of services provided by human capital. Even the technical side of the input of human capital is difficult to assess, in spite of the studies that indicate some success in using "peer review" methods to assess quality. This problem is compounded at the output end where quality becomes a multifaceted variable.¹ One suggestion is that quality variation be measured by the price component of inputs and output while the usual procedure is to assume constant quality over the range of productivity increases. But the first alternative is not applicable to the health sector given the nature of price determination where competition is minimal and cross subsidization is used to compensate for high cost procedures while the second suggestion is dubious given the rate of change of medical science and technology. The problem of accounting for quality often comes down to a subjective consideration of the likely direction of change owing to productivity increases. Thus while examining the apparent productivity changes observed through input substitution the researcher must control for quality or at least remain conscious of quality factors.

Another caveat deals with the paradoxical situation of recognizing productivity as an intermediate goal. That is, if health improvements are the ultimate objective of the productive effort, then increased productivity by one group of manpower; i.e., geriatricians, may appear to result in decreased productivity by another group, for example, cardiologists. The

1. For an Ontario study of peer review see K. F. Clute, The General Practitioner, Toronto: University of Toronto Press, 1963).

example here is that recent advances in medical practice have often resulted in postponing death rather than curing or preventing illness. This situation involves apparent productivity shifts among manpower categories.

A further complication is the close connection between factors influencing economies of scale and those influencing input substitution possibilities. There were, for instance, only 60 electroencephalography technicians and 150 electrocardiography technicians in Ontario in 1968. This may be because these less trained labourers are not being sufficiently substituted for more highly trained professionals. But it is more likely that the capital equipment required to employ these individuals is not available in all hospitals because the demand for their services is not sufficient to justify the cost, and also that the training of these individuals is on such a small scale that it is very expensive to provide a central training institution. Thus, it is clear that possibilities for input substitution in a regionally dispersed country like Canada is highly dependent on economies of scale.

III Potential for Productivity Growth

Sources of Past Productivity Growth

One could probably attribute a fairly large portion of the growth in postwar medical productivity to two general sources. Since the 1930's there has been a drastic reallocation of physicians' time from house calls to office and hospital visits and to telephone consultations. In the U.S.A. in 1930 home visits comprised 45 percent of all outpatient visits; in 1965

they were only 4 percent of the total.¹ Since the average house call takes up to three times as much physician time as an office visit, it is apparent that this substitution of patient travelling effort for that of the physician has resulted in substantial time saving for the physician. It is also apparent that this particular source of productivity growth is unlikely to contribute much in the future to productivity increase unless for example there are further innovations related to significant developments in telephone communications such as would perhaps occur if video machines were widely available as attachments to telephones.

The second major source of productivity growth can be traced to advances in bio-medical knowledge and techniques, and in the development of antibiotics.² During the late 1940's and early 1950's many drugs were introduced which were highly effective against influenza, pneumonia, tuberculosis, and other infectious diseases which had large impacts on morbidity and mortality. These drugs included penicillin, streptomycin, chlortetracycline, the first broad spectrum antibiotic in 1948, and tetracycline in 1953. Since 1956, however, the tenor of technological development has changed. New advances have typically been of the kind which have only a small impact on general health and, at the same time, are highly resource intensive. Radiation therapy, open heart surgery and

1. U. E. Reinhardt, "An Economic Analysis of Physicians Services", (unpublished Ph.D. dissertation, Yale University, 1970) p. 13.

2. V. Fuchs & M. Kramer, "The Market for Physicians' Services in the United States, 1948-1968", (New York: National Bureau of Economic Research, unpublished working paper) Ch. 17.

organ transplants are examples of this.

There has also been a fairly steady substitution of ancillaries for physicians over the last two decades. For example, Weiss found that the 1950 input coefficients for health manpower would have meant that by 1960 46,000 more physicians in excess of the actual total number would have been employed and there would have been far fewer employees in the job requiring less intensive medical training.¹

Barriers to Substitution

The general production function concept, even given the above noted difficulties of analysis specific to the health sector, allows for a good deal of flexibility. With regard to input substitution there are three areas in which technical and economic feasibility have been reported. From an activity analysis point of view, input substitutability exists where:

a) certain functions can be reassigned to personnel with lesser training than those currently responsible for their execution,

b) certain functions can be transferred from the provider to the consumer of services,

c) certain functions can be performed by less manpower and more capital.²

-
1. Reinhardt (1970) p. 11 and J. Weiss, "The Changing Structure of Health Manpower" (unpublished Ph.D. dissertation, Harvard University, 1966).
 2. E. Ginzberg, Men, Money and Medicine (New York: Columbia University Press, 1969) p. 134.

Our principal interest is presently in the first two of these cases in which substitution is strictly in terms of the labour input. Capital substitution and technological advancement will be considered in a latter part.

Before entering into a review of the manpower substitution literature it should be understood that the flexibility of the production function is limited by structural factors of the health sector. These include licensure procedures and traditions, reimbursement mechanisms, and attitudes of health professionals and consumers. As will be seen shortly it is these obstacles to innovative substitution that have limited productive efficiency in Canada. They are dealt with at this point to put the empirical investigations reviewed later in the perspective of having most often accepted these boundaries to substitution as unalterable.

The impact of the current system of registration, certification and licensing can be reviewed within a cost-benefit framework but first it should be pointed out that there are fine lines separating each of these methods of regulation.¹ Theoretically, registration is available to anyone choosing to produce a given service irrespective of qualification, while certification is restricted to individuals possessing a certain skill level, though anyone is allowed to provide the service and licensing is a mandatory requirement limited to producers who demonstrate a certain competence. In practice, however, there is considerable confusion of these legal constraints such as the licensing of physiotherapists, radiological technicians and

1. R. D. Fraser, "The Economic Consequences of Licensing", Mimeograph, Queen's University, 1973.

the recommended licensing of dental technicians who receive less than four years of post secondary education, while clinical psychologists receiving 8 to 9 years of advanced training are merely registered. Also, the wide range of training, from 2 to 4 years, for registered nurses reduces the credibility of regulation, while confusing registration with what is, in fact, certification. It may be more justified, in practice, to view health care personnel with respect to their varying degrees of monopolistic power which is the source of most of their control and is, in any event, indirectly attributable to specific legal considerations.

In general, the licensure system benefits are:

- a) a guarantee of minimum competence at some point in the professional's career,
- b) increased visibility and supervision of licensees, and
- c) reduction of choice difficulties experienced by consumers

unaware of differentiating criteria.

These all can be reduced to a concern for setting out a guarantee of minimum standards of care. However, the success of this method of guarantee has not generally been supported by empirical evidence.¹

The potential benefits seem especially low when viewed against the background of the costs associated with licensing. First, the monopolistic power implied by licensure probably raises the price of the manpower input. The same power probably restricts the supply of manpower and thus reduces opportunities for substitution. Secondly, the discretionary power of these monopolists can be used to the detriment of particular segments of the population. For instance, the female representation in top levels of

1. Klute, Op. cit.

health care personnel is restricted as is the representation from lower socio-economic classes which would expand the numbers of prospective candidates. Thirdly, the reluctance to introduce new methods that would increase competitiveness in the profession is the result of vested interest in the benefits of monopoly. Finally, and most important, the effect of licensure and corresponding regulations is to restrict the number of tasks that can be provided by other personnel of lesser training. Examples of this trend are the replacement of midwives by obstetricians and specialists for general practitioners. In addition, the functions that could, in fact, be performed by auxiliaries, independent of professionals, are performed under supervision because of malpractice threats created by licensure responsibility. The supervision in these cases is often technically superfluous and economically expensive.

Monopoly power aside, the impact of licensing is still significant. By narrowly defining the job functions of each level of auxiliary worker, the system reduces job mobility and career income potential such that many lower level categories are considered dead end jobs.¹ Although there are certainly some individuals available and willing to accept this kind of career, and dead endedness is not necessarily a negative career characteristic, there are clear limits to the quality of personnel attracted and probably side effects of high rates of personnel dissatisfaction and turnover. The fact that females predominate in many categories of allied workers along with the high rate of turnover and lack of prestige among these groups has probably had a significant impact on policy makers. That is, if the less comprehensively trained workers

1. In all fairness it should be noted that nursing has been an occupation in which upward social and economic mobility has been facilitated perhaps largely because of the team work involved in health care facilities such as the hospital, where such teams include a variety of occupations which all but span the range of socio-economic class.

are to be allowed greater responsibility, they first must demonstrate greater potential for dependability and permanence. Of course, this is a vicious circle of low wages failing to attract quality personnel and low quality personnel failing to provide justification for investment in human capital, increased responsibility or incomes. At the centre of the circle is the restrictive licensure system. A further impediment to substitutability is the highly localized nature of licensure resulting in expensive retraining of personnel to meet standards and a major cause for career dropouts. If the standardization of personnel could be accomplished by some nationally accepted method this in itself would advance the value of allied workers.

With regard to remuneration mechanisms, Dr. Dollery has pointed out that the government control of the purse strings in Ontario is only the first step towards influencing the manpower utilization patterns.¹ He believes that to cause new types and great numbers of allied health workers to be employed will require a breakdown of the fee-for-service payment system and the establishment of new reimbursement mechanisms. This fact explains a good deal of the disparity of utilization patterns between Canada, the United States and Great Britain. In the U. S., there are economic profits available to entrepreneurial health professionals who employ allied health workers on a salary basis but are paid for their work on a fee for service basis. In Great Britain, the conversion from fee for service to salaried health professionals has allowed for central planning of substitution programs. In Ontario, the health professions covered under O.H.I.P. have very little incentive to hire auxiliaries because they can collect fees only for the

1. C. T. Dollery, "Delivery of Care. Acute and Chronic Illness", The Greater Medical Profession (Josiah Hacy Jr. Foundation, etc., see full description above) p. 78.

services that they perform themselves. Nor has the government yet taken the initiative to provide a system under which the use of auxiliaries is mandatory. There are a number of methods for providing economic incentives to practitioners such as subsidizing the salaries of auxiliaries, or payment to practitioners on a capitation basis. Suffice to say that unless the incentives are made large they will not be sufficient to overcome the managerial disutilities associated with the employment of auxiliaries.

The final barrier to substitution that we will note is the attitudes of professionals, consumers, and the current allied health workers. This is a largely unproven hypothesis given the many-times reported acceptability of new types of health workers by all three groups.¹ On the other hand, this acceptance may come as the consequence of a well-educated study group. In general, if misconceptions are a major source of aversion to auxiliary workers, it is clearly necessary that the educational program required to change these perceptions be instituted. As long as this provision of information is relatively cheap, it is not likely to alter the economic feasibility of substitution.

Greater Involvement of the Patient

The notion of patient participation in the delivery of health services is being very broadly defined. By patient participation is meant

1. Most of the studies surveyed in the following section included acceptability statistics. The Yankaver, et. al. survey of pediatric practices revealed the percentage of pediatricians in favour of task delegation for a large number of office functions. See A. Yankaver, J. Connelly, and J. Feldman, "physician Productivity in the Delivery of Ambulatory Care: Some Findings from a Survey of Paediatricians," Medical Care, January-February 1970), p. 827.

A survey of acceptance is also reported in M. C. Collins and C.C. Bonnyman, "Physician Assistant and Nurse Associate: A Review", The Institute for the Study of Health and Society, Contract No. HSM 110-70-371, Washington, D.C.

not only direct involvement in the production of psychomotor tasks by patients but also activity in choosing levels of utilization, phasing treatments, monitoring and sustaining therapy, and executing health maintenance. The issue posed is whether there exist opportunities for increasing the role of the patient in providing health care which is cost effective. It is widely presumed that such opportunities do exist and a rather massive research literature has developed which is concerned with this issue. However, most of the research is fragmentary in the sense that it addresses only a subset of the relevant issues. In particular, the literature rarely resolves the question of economic feasibility.

In order to focus attention upon health status and behavioural outcomes rather than pedagogy, it is useful to review the literature beginning at the outcome level. This organization also conforms to the research strategy; e.g., first determine whether an outcome is desirable, then examine methods of achieving it. The survey below is far from complete.

As noted above, the notion of health status has no universally accepted operational definition. The literature on the impact of health care on health status is consequently not well unified. Extensive biomedical and clinical research supports the view that specific programs of prevention or treatment appear to suppress or remedy specific complaints. There are very few formal attempts to assess the economic efficiency of alternative modes of care. The rather superficial literature on the Health Maintenance Organization concept suggests that preventive care and early diagnosis are preferable to episodic care. These studies were

based on aggregative outcome data and provide minimal insight into the total care received or the processes of cost reduction.

Although there is no clear agreement as to which health services should be produced in greater quantities in order to improve health status, several researchers have sought to evaluate particular procedures. These investigations may be grouped under the headings of 1) health maintenance activities, 2) compliance with medical treatment, 3) utilization of the health care system, 4) patient self-care and 5) the effects of patient knowledge.

Preventive care studies have examined the usefulness of periodic examinations. Fromco et al., in a panel study of salaried employees of a New York firm found that major disease was discovered over a ten year period in asymptomatic form in 59 percent of those examined.¹ David, however, questions the contribution of such examinations to the reduction of disability and mortality.² Wode et al., concluded from analyses of medical records for a group of 765 employed males that disabling diseases were diagnosed frequently in the course of periodic examinations. Grinaldi in an analysis of match groups of employees of General Electric established that the costs of medical care for the group not being examined exceeded that of the experimental group by more than the cost of the periodic

-
1. S. C. Franco et al., "Periodic Health Examinations: A Long Term Study 1949-1959", Journal of Occupational Medicine, Vol. 3, No. 1 (1961).
 2. W. D. David, "The Usefulness of Periodic Examinations" Archives of Environmental Health, Vol. 2, No. 3 (1961).

examinations.¹

The effectiveness of the health care system in improving health status is widely thought to be eroded by non-compliant patients. Considerable research has focused on the ignorance of the significance of instructions or the underlying rationale. Curtis and Schwartz investigated the compliance of elderly, ambulatory patients with programs of chemotherapy and found that more than sixty percent of the patients were non-compliant.² The patients omitted medications, relied on self-medication and obtained incorrect dosages. Similar results regarding the extent of noncompliance were found in studies by Pragoff (1962), Preston and Miller (1964) and Davis (1966).

Some of the best work on patient participation has examined the potential for increasing patient responsibility for his care. The need for continuous management of diabetes, chronic heart disease and hypertension in particular has suggested that it would be cost-effective to transfer some responsibilities to the patient. A study of an education program for heart patients revealed that the study group required approximately one-third as many hospital admissions and one-fourth as many days of

-
1. J. V. Grinaldi, "The Worth of Occupational Health Programs: A New Evaluation of Periodic Physical Examinations", Journal of Occupational Medicine, Vol. 7, No. 8 (1965).
 2. E. B. Curtis, "Medication Errors Made By Patients", Nursing Outlook, Vol. 9, No. 5 (1961), D. Schwartz, et al., "Medication Errors Made by Elderly Patients", American Journal of Public Health, Vol. 52 (1962).

hospital care. A study using parents for a good part of inpatient care demonstrated a reduction of hospitalization by 40 percent and reduced direct personnel costs went from \$10.31 per day to \$3.81.¹

The literature investigating the extent of patient knowledge indicates that even for common and highly publicized problems, public beliefs and attitudes are inconsistent with a preventive mode of care and early detection. Although this problem is most profound among lower socio-economic and educational strata, it is widespread. It is not clear, however, if one can expect desirable changes in patient behaviour to result from increases in knowledge.

Major Policy Decisions

The above discussion of barriers to substitution points out areas in which research is needed and policies chosen that will allow access to potential productivity gains. The emphasis here is on co-ordinated progress towards substitution possibilities such that the required policy changes will be forthcoming when substitutability has been proven to the satisfaction of the powers that be and comprehensive programs for substitution drafted. Note that decisions to work within the narrow confines of current policies will severely limit the productivity gains.

The policies referred to are those on licensing procedures, reimbursement mechanisms and educational programs. A long hard look at

1. S. G. Rosenberg, "Patient Education heads to Better Care for Heart Patients", HSMHA Health Reports, Vol. 86 (1971).

and estimation of the economic consequences of licensure in Ontario's health sector is long overdue. Even if the general concept of some form of licensure is accepted there is a need for overhauling the status of each of the job categories. As for reimbursement methods; there will be few productivity gains if health professionals cannot be given personal economic incentives to increase their utilization of auxiliaries. There is a pressing need for new directives in this area due to the ever present possibilities for substitution that are not being taken advantage of. Finally that rapid change in the methods of health care delivery will not be acceptable to the public without extensive educational programs to re-establish perceptions of health care workers should be anticipated along with programs to unify the health care team.

IV Literature Survey

Primary Care

Research on the possibilities for manpower substitution are usually concerned with demonstrating the capability of less expensive, less comprehensively trained personnel, to perform the functions of more expensive personnel. The manpower categories for which this hypothesis has been tested are: general practitioners, pediatricians, obstetricians, ophthalmologists, chronic care professionals, internists, gynecologists, and dentists. In reviewing the results of these studies, it is important to note that the researchers have been, to date, more interested in testing the methodology of analysis and the general hypothesis of substitutability rather than explicitly establishing

the degree of substitution possible in each specialty. On the other hand, the time has come for greater organization of the effort to rationalize manpower productivity. This can best be accomplished by first evaluating the methodologies of production analysis and secondly applying the best method to each case where substitution is likely.¹ For our purposes, the methodology will be noted for each study. In general, the approaches have been either production function estimates, activity or time and motion descriptions, or utilization questionnaires.

The most often analysed case is that of general practitioners. Reinhardt (1972) and Reinhardt and Yett (1970) use cross-section data from two large questionnaire surveys to estimate a production function for general practitioners. They only reported the results for general practitioners because they found them very similar to estimates for obstetricians, gynecologists and internists.² and ³ Including as inputs measures of R.N.'s, technicians and office aids per physician, Reinhardt (1972) estimated that the marginal productivity of an auxiliary tended to increase up to the level of about one aid per physician, diminishing thereafter, and becoming negative at a level between 5.0 and 5.5 aids per

-
1. U. E. Reinhardt, op. cit. (1973) for a discussion of productivity analysis methodology.
 2. U. E. Reinhardt, "A Production Function for Physician Services", The Review of Economics and Statistics, Vol. 54 (February, 1972) p. 61.
 3. U. E. Reinhardt and D. A. Yett, "Physician Production Functions Under Varying Practice Arrangements", Technical Paper Series No. 1, Rockville, Md., Community Profile Data Center (taken from U. E. Reinhardt above).

physician. The general results of these two studies were that at prevailing wages, the average solo practitioner could profitably employ almost twice as many ancillary personnel as are currently being used; i.e., from 1.9 to 4 aids per physician. This would result in approximately a 28 percent increase in the number of patient visits per week per physician.

Kehrer and Laretsky (1972), using different data and different functional specifications,¹ found results surprisingly similar to Reinhardt's. They concluded that, in general, a doubling of allied health personnel input would increase total patient visits by 20 to 25 percent. The marginal product curves for ancillaries tended to become negative at generally higher levels of aids per physician than in Reinhardt's work, and there were wide variations across medical specialties. Their estimates implied that internal medicine and general practice appeared most capable of absorbing additional ancillaries. This is consistent with the general feeling that the less esoteric the specialty, the greater the opportunities for delegating tasks to paramedicals. However, this behavior may also reflect the differential constraints within which delegation of tasks is possible. Kimbell and Loront (1972) estimated production functions for the private practitioners. Although their actual estimates differ somewhat from previous work, their general conclusions are highly consistent with the results of previous studies.

Activity analysis was used by Golladay, Miller and Smith (1971, 1972) to develop a normative model of how a primary care practice should be

1. B.H. Kehrer and H.W. Zaretsky, "A Preliminary Analysis of the Productivity of Allied Health Personnel in Primary Medical Practice", (unpublished working paper, Center for Health Services Research and Development, A.M.A. Chicago 1972, p. 98. (referred to in J. Hadley, "Factors Affecting Health Manpower Productivity", (March, 1972) Economic Analysis Branch, Social & Economic Analysis Division, National Center for Health Services Research & Development.)

organized.¹ The procedure consisted of enumerating 263 tasks in eight major categories of activity which fully described a primary care practice. In order to determine the frequency with which tasks were delegated, five medical school students acted as observers in a sample of primary care practices in Wisconsin, Vermont, and North Carolina over a two-week period. In the actual analysis, subsets of tasks were aggregated to form medical services.

Solution of their model provides the minimum labour cost and input mix for any given level of output. Their results were predicated on the assumptions that the physician devotes 28 hours a week on direct patient care, that all tasks not classified as being part of some medical service can be delegated to the same extent as classified tasks and that output is at a level such that further delegation of tasks is not possible. Given these assumptions, they find that productivity of the individual physician could be increased by 74 percent (from 147 visits to 265) by using a physicians' assistant. It should be emphasized, however, that these are potential gains which depend crucially on the above assumptions and the additional implicit assumption that physicians are efficient in organizing their practices.

With regard to physicians' assistants, it is also interesting to note that another study, Scheffler and Stinson (1972) estimated that the mean

1. F. Golladay, K. Smith and M. Miller, "Allied Health Manpower Strategies: Estimates of the Potential Gains From Efficient Task Delegation", Health Economics Research Center, University of Wisconsin Research Report Series No. 15, November 1971 and Journal of Human Resources, Vol.VII, Spring 1972, p. 225.

starting salary of a P.A. to be \$10,501, while the possible gain to the physician is two to three times as much. In his analysis of the capabilities of nurse practitioners, Kaku¹ compared the diagnostic records of doctors and nurse practitioners for 1000 patients undergoing periodic examination. In 70.3 percent of the examinations there was complete concurrence of findings. Fourteen percent of the physicals revealed greater detail of observation by nurses while physicians noted and nurses apparently overlooked 5 percent of positive physical findings. All of the findings that were recorded unilaterally were of minor significance.

Pondy (1970) with "before" and "after" time motion studies has shown a significant amount of time saved after a Duke University trained physician's assistant joined a private practice. The assistant saved the physician 43 percent of time formerly spent in routine tasks. This time saved and the extra help from the assistant enabled the pair to devote 16 percent more time to the average patient and 36 percent more time to significant new illnesses. There was also increased output of preventive care in that 13 times as many work-ups were carried out by the team than by the physician alone.² Lewis (1967) using a descriptive activity analysis reported on the risk-taking patterns of nurse practitioners, the processes of care, acceptance by patients, outcomes of care and cost

-
1. K. Kaku, F. Gilbert, R. Sachs, "Comparison of Health Appraisals by Physicians Assistants", Public Health Reports, Vol. 88 (1970) pp. 1042-46.
 2. L. R. Pondy, "Physician's Assistant Productivity", Ayden, North Carolina. Unpublished, Duke University, Durham, North Carolina, April 1970 from "Physician Assistant, Nurse Associate, A Review", C. Collins & G. Bonnyman, January, 1971.

of care from medical clinics and nurse practitioners respectively. The results were that a significant shift in the preference for nurses as performers of almost all nurse medical functions was noted. While there was no difference in progression of disease or in death rates between control and experimental groups, at the end of the year of observation patients treated by nurse practitioners showed significant reductions in disability and frequency of complaints.¹

The evidence presented above points out, above all, that considerable substitutability exists for primary care practices. In addition, the efficacy of personnel trained specifically to assist physicians appears to be strongly supported as well as the hypothesis that less comprehensively trained manpower can be substituted for physician input.

Pediatrics

In keeping with the general concept that the less "esoteric" specialties are more open to manpower substitution in present circumstances the next most observed case is that of pediatricians. Using surveys of ancillary utilization patterns and the distribution of tasks delegated by pediatricians, Yankauer, Connelly and Feldman (1970) reveal some interesting facts.² They conclude

-
1. Lewis & Resnick, "Nurse Clinics & Progressive Ambulatory Care", New England Journal of Medicine, Vol. 297 (1967) pp. 1236-41.
 2. A. Yankauer, J. Connelly, J. Feldman, "Physician Productivity in the Delivery of Ambulatory Care: Some Findings from a Survey of Pediatricians", Medical Care (January-February, 1970) p. 827

that, "the tasks least frequently delegated are those which involve clinical judgment or a more intimate patient-caretaker relationship", while, "the patient care tasks most frequently delegated are those which reflect the most urgent pressures of practice (telephone calls) or those most easily routinized (interpretation of instructions and routine history-taking)", they also suggest that pediatricians in northeastern U.S.A. delegate less frequently and have fewer ancillaries than those in the West. This latter statement agrees with conclusions by Rafferty (1971)¹ Riddick (1971)² and Hadley (1973)³ that physicians are more likely to utilize ancillaries the greater the demands on their time.

Ott, Machotka and associates have in press a report comparing the performance of nine child health associate graduates with that of freshmen medical students and pediatric residents with respect to the pediatric portion of the National Board of Examiners. The associates scored higher on the basic science than either the residents or the medical students while the associates averaged 83 percent on the clinical pediatrics section compared to an average score of 94 percent for medical students.⁴

-
1. J. A. Rafferty, "Patterns of Hospital Use: An Analysis of Short Run Variations", Journal of Political Economy (January-February, 1971).
 2. F. A. Riddick, et al., "Use of Allied Health Professionals in Internists Offices", Archives of Internal Medicine (May 1971) pp. 924-931.
 3. J. Hadley, "Physicians Incentives for Joining Group Practices", Economic Analysis Brancy B.H.E.W. HSMHA, NCHSRD., 1973.
 4. P. Matchotka, et al., "Competence of Child Health Associates: A Comparison of Their Basic Science & Clinical Pediatric Knowledge...", American Journal of Diseases of Children (in press) taken from C. E. Lewis, op. cit.

Bergman (1966) using a time and motion study has shown that 48 percent of the pediatrician's time is spent in patient care, 12.5 percent on the phone and nine percent on paper work. Of the time spent in patient care, 50 percent is in well child care and another 22 percent is with children with minor upper respiratory infections. In actual practice, pediatric nurse associates are able to do most of these types of pediatric care.¹ Silver's group in Colorado has shown that in attendance only one-half day each week, the pediatric nurse associate was able to manage 82 percent of the 2,735 visits studied. She managed 71 percent alone and required consultation in 11 percent with a physician.²

Other Areas of Health Care

The research carried out with respect to obstetricians, gynecologists, internists, and optometrists use similar approaches and come up with similar results.^{3, 4, and 5} First, there are significant possibilities for productivity

1. A. Bergman, S. Dassel, R. Wedgewood, "Time-Motion Study of Practicing Pediatricians", Pediatrics Vol. 38 (August, 1966) pp. 254-63.
2. H. Silver, L. Ford, S. Stearly, "Program to increase health care for children: pediatric nurse practitioners", Pediatrics, Vol. 39 (1967) pp. 756-60.
3. J. C. Record and H. R. Cohen, "The Introduction of Midwifery in a Prepaid Group Practice", American Journal of Public Health Vol. 62, No. 3, (1972) pp. 354-60.
4. I. Schwartz, "Vision Acuity: The Professional Versus the Non Professional", from C. E. Lewis Evaluating the Performance..., op. cit., p. 98.
5. U. E. Reinhardt, op. cit., (11).

increases through the employment of trained specialists' assistants as well as standard auxiliary personnel. Secondly, neither general practitioners nor specialists have sufficiently exploited this source of productivity gain. Thirdly, the general trend is for more pressure on physician time to elicit greater use of auxiliaries.

Because all of the above-mentioned studies were done in the U.S., we might question their applicability to Canadian conditions. In general, there does not appear to be any grave differences that would discount the American research applicability here. There has been little in the way of empirical measurement to confirm this although the concepts have been well accepted by many interested groups. The college of Family Physicians of Canada firmly supported the introduction of a nurse practitioner in 1971.* In the 1970 study for the Committee on the Healing Arts, Fraser concluded that the public health nurse would add significantly to the quality of care received in rural and low income areas where general practitioners are not available.¹ He also concluded that, "there is ample justification, if not for the training and use of midwives in Ontario, at least for broad experimentation in their use in hospitals, group practices and public health centres",²

Dentists are far easier to analyze with respect to auxiliary workers because of the limited job functions and already well-divided labour force.

1. R. D. Fraser, Selected Economic Aspects of the Health Care Sector in Ontario (Toronto: Queen's Printer, 1970)p. 188.

2. ibid., p. 186.

The dental hygienist, dental assistant and dental technician all have specific functions that are easily identified. There is also a firm segment of literature supporting increased substitution of auxiliary personnel in dental offices. In the Royal Commission Report on Health Services (1964), McFarlane concludes from a survey of Dental practices in 1958 that dentists on average could double the number of auxiliary employees hired from just over one to three.¹ This same kind of proof for substitutability is continually being produced in the U.S. For instance, the Philadelphia Department of Public Health dental program reported finding that three techno-therapists under the supervision of a dentist and working with two assistants and a clerk can provide as much total care as four dentists each working with one dental assistant.² And yet, in Ontario, in 1968 there were 3,000 dentists employed along with 250 dental hygienists, 600 dental technicians and 4,000 dental assistants. This ratio of auxiliaries to professionals would indicate that dentists persist in maintaining inefficient operations.³

1. B. A. McFarlane Dental Manpower in Canada, Royal Commission on Health Services 1964, Ch. 6.
2. For a complete discussion of economic aspects of dental care see P. Feldstein "Financing Dental Care: An Economic Analysis", under NCHSRD Grant # HS 000 41.
3. For more recent studies of the potential for dental auxiliaries, see the following: Hord, A. Bruce, and R. L. Ellis, "O.D.A. Demonstration Project Phase One: A Training Program of Expanded Functions for Dental Auxiliaries," Journal of the Ontario Dental Association, (Nov., 1972), pp. 306-310. Romcke, R. G., and D. W. Lewis, "Use of Expanded Function Dental Hygienists in Prince Edward Island Dental Manpower Study," Journal of the Canadian Dental Association, (April, 1973), pp. 247-262. Ellis, R. L., et al., "Ontario Dental Association Demonstration Project Phase Two," Ontario Dentist, (Feb., 1974), pp. 15-18. Hord, A. B., et al., "The Ontario Dental Association Demonstration Project on Dental Auxiliaries with Expanded Duties," Ontario Dentist, (June, 1974), pp. 14-18.

The discussion of manpower substitution comes down to the simple decision by practitioners not to rationalize their productive effort. Reasons often given for this general reluctance are: fear of quality of care deterioration, fear of patient resistance to auxiliaries, loss of physician-patient personal relationship, managerial disutilities, increased likelihood of malpractice suits, costs of on-the-job training, licensure limits on what tasks ancillaries are qualified to perform¹, and the effect on physician's perceptions of the particular educational programme by which they are trained. These items should be the centre of further research given that the economic feasibility of substitution has been shown but not adopted by the majority of practitioners.

Capital Substitution

Possibilities for increasing health manpower productivity through the substitution of capital for other production function inputs have not received much direct attention by economists² and this is not the principal aim of this present section. There is, however, a fairly large literature developed by engineering and computer oriented specialists. This literature naturally reflects the perspectives and concerns of those groups rather than those of economists. The main issue in most cases is whether the new type of capital (or new technology embodied in some type of equipment) increases output or reduces costs per procedure relative to the prior modality. In other words, the usual objective has been technical feasibility demonstration rather than economic.

-
1. R. Fern, "Economic Consequences of Accepting the Concept of the Greater Medical Professions", The Greater Medical Profession (New York, 1973) p.168.
 2. Work on economies of scale of different health care facilities of course implicitly deals with capital-labour substitution.

V A Research Agenda

We have drawn attention in the above discussion to several areas in which potentially valuable research could be carried out in order to facilitate the development of a more efficient health care delivery system through manpower substitution. We do not propose to go over these in detail. Rather, we propose a two pronged thrust for research effort along the following lines.

First, a comprehensive matrix of studies should be done on each of the many distinct health care occupations to determine the nature of the tasks they perform. The set of abilities required to perform these tasks that could be gained through experience or provided through educational institutions, the actual set of abilities now acquired in their educational programmes, and an assessment of the effectiveness of the legislation under which they are now registered, certified or licenced, should all be investigated.

The second major thrust involves carrying out a matrix of highly applied cost benefit studies on each of a selected set of potentially significant substitution possibilities. Included in this set would be the substitution of dental health nurses for dentists, midwives for obstetricians, public health nurses for general practitioners, denturists for dentists, chiropractors for orthopaedic surgeons, to name but a few.

CHAPTER VIII

Advances in Health Care, Science and Technology

I. INTRODUCTION

As we have seen, the health care industry must now be counted as one of the principal sectors of the Canadian and Ontario economies. In addition, we have noted that the recently acquired responsibilities of government in the field of hospital and medical care insurance have made government one of the major elements of the structural organization of this industry. We now wish to consider that part of the health care industry that is engaged in the work of advancing the frontiers of medical science and technology. Though often ignored in what frequently appears to be a myopic search for short-run allocational efficiency and distributional equity,¹ the dynamic aspects of the health care industry are of course vital to the present and future health of the industry itself.

Our concern for the processes by which advances are made in the science and technology of delivering health care and for the evaluation of such advances, or more simply for the economics of health research, stems in part from the very significance of the role of government in this area and thereby, the almost complete absence of market forces as potential agents for ensuring efficient resource allocation. While the entry of government into the health care delivery field as the principal funder of health care is of recent origin, its role in the funding of health research is not. For some time, its role in the health research field has been major and at the present time can be described as pre-eminent. An overwhelming percentage of the general financial support of research now comes from the provincial and federal treasuries either directly from research funding agencies or indirectly through the budgets for higher education and hospital care. In

1. Notwithstanding the fact that in the long run we're all dead!

addition, contributions to the overall research effort are made by the many government operated research laboratories.

This past role of government in the funding of health research is of course related to the very nature of research in general and health research in particular. Research activity is characterized by immense levels of uncertainty; there is almost a complete lack of knowledge of the processes by which specific research resources are transformed into valuable research output. In addition, the value of the benefits of successful research ventures, even if such benefits are recognized during the lifetime of the researcher in question and can be evaluated, are seldom appropriable by the researcher. Finally, once obtained the social costs of using new knowledge are usually very low relative to the costs of the research itself; these are largely the costs of transferring the new knowledge from the researcher to the potential users. This last characteristic of indivisibility together with the preceding ones of inappropriability and uncertainty seem to justify the significant role of government in the funding of health research; they certainly portend the failure of the market as the chief mechanism for allocating resources to and within this sector.

Our second source of concern is with the relationship that exists or should exist between the industry involved in the production of goods and services and the research and development effort that supports it. Along these lines, there would seem to be, in addition to the fundamental justification advanced above for the role of government in the health research sector, the more pragmatic one which is associated with its responsibilities for the delivery of health care. As the principal funder of health care and thereby

as the entity bearing a principal responsibility for the overall organization of the health care delivery system, government would seem to have an extra incentive to insure that the research and development effort supporting this delivery system is adequate. Just as a profitable, privately-run industry would ensure that the level of its overall research effort was "reasonable" and that such effort was being sensibly deployed according to some set of established priorities, so also might we expect government to act in the face of its responsibilities in the entire health care industry.

In view of these concerns, a review and evaluation of the health research sectors in Ontario and Canada seems justified. And indeed, the Ontario Council of Health has recently published one such study.¹ A one sentence summary of the results of this document is that the research effort supporting the health care delivery system is heavily unbalanced, fragmented, largely unorganized, and seemingly too small by half.² Taking the last point for example, expenditures on health research and development in 1971 in Ontario and Canada amounted to the equivalent of only 2.2 percent of total expenditures on health care. In contrast, some 5 to 10 percent of total sales revenue is allocated to research and development in industries with rapidly developing technology and which are closely related to the health sector. Similarly, all manufacturing industries taken together devote on average over 4 percent of total sales revenues to research and development as judged by countries for which satisfactory data are available.

1. R.D. Fraser, "The Economics of Health Research," pp. 26-202 in Report of Committee on Health Research (Toronto: Ontario Council of Health, 1974.)

2. See Fraser, op. cit., especially Chapter II, pp. 50-87 and Chapter VI, pp. 147-166.

In this section we propose to set forth a set of priorities for "research on research." In so doing we are conscious of the limited knowledge base from which we start. There are but a few studies that can be considered as contributions in this area. These few will be briefly surveyed before we begin our attempt to set out a blueprint for research.

In starting out on our task, we have taken courage from a recent paper of E. Mansfield.¹ Mansfield had been asked to write a paper for Science on the contributions of research and development to economic growth in the United States. By implication, his judgment would appear to be that our understanding of the general processes of research and development in the economy at large is very small indeed, since instead of speaking directly to the proposed subject this paper is itself almost wholly a blueprint of the many aspects of the research and development process which remain yet unresearched.

II. LITERATURE SURVEY

Z. Griliches

The seminal work in the estimation of the returns to research effort is that by Griliches on the outcome of hybrid corn research.² His calculations were based on assumptions of research costs, the prices of seed corn and corn as a consumption good, elasticities of demand and supply,

1. Edwin Mansfield, "Contribution of Research and Development to Economic Growth in the United States," Science, Vol. 175 (February 4, 1972), pp. 477-86.

2. These results were reported in Z. Griliches, "Hybrid Corn: An Exploration in the Economics of Technological Change," Econometrica, Vol. 25 (October, 1957), pp. 501-522; and _____, "Research Costs and Social Returns: Hybrid Corn and Related Innovations," Journal of Political Economy, Vol. 66 (October, 1958), pp. 419-31.

etc., that would cause his final estimate of the rate of return to be an underestimate of the likely rate. For example, in estimating the sum of private and public expenditures on hybrid corn research over the period from 1910-55 from data principally collected by the U.S. Department of Agriculture and private seed companies, he used overall figures on "corn production research: agronomic phases." Such figures included funds allocated to corn research other than hybrid research. Further, he assumed that continuing maintenance research expenditures of some three million a year from private and public sources would be required but that there would be no further increases in the productivity of hybrid corn as a result of these expenditures. Also assumed is the absence of any spillover effects between hybrid corn research and hybrid sorghum research for example.

In a similar way, Griliches assumed a price of hybrid seed corn that was fifty percent higher than the one prevailing at the time of his research; he assumed that the superiority of the hybrid corn over its counterpart, open pollinated varieties, was only fifteen percent even though estimates of its superiority by breeders actually ranged from fifteen to twenty percent; the elasticity of supply for corn was chosen so as to yield a minimum estimate of the benefits accruing as a result of the introduction of the hybrid corn; and, finally, the price of the final product corn and the number of acres seeded with it were likewise underestimated.

The resulting estimates of the return to hybrid corn research, though by design underestimates, are clearly rather startlingly large. "At least 700 percent per year was being earned, as of 1955, on the average dollar invested in hybrid corn research."¹ Even if all past returns were ignored, this figure would only fall to some 400 percent. When the estimates

1. Z. Griliches, "Research Costs and Social Returns...", p. 419.

were made according to the internal rate of return procedure, an internal rate of return of some 35 to 40 percent was found; and similarly, the benefit cost ratio was estimated to be 150. When the discount rate of 10 percent was used instead of the 5 percent rate, the rate of return estimated fell to 689 percent and the benefit-cost ratio fell to 70.

On reading the results obtained by Griliches, one is immediately led to enquire if the hybrid corn breakthrough is an example of one success among many failures. Limited study of the overall return to research on agricultural products has, however, been completed; the results of it leave one with the impression that the hybrid corn research was indeed very successful but that research in this area has on average been successful. For example, T. W. Schultz¹ had assembled data on the overall returns to research on agricultural products that yield the estimate of a rate of return of 171 percent using a 5 percent discount rate. In turn, these estimates can be compared to the 100 to 200 percent rate of return to all expenditures on research and development in the entire U.S. economy as presented by B. R. H. Ewell.²

This work by Griliches is seminal in laying out the general methodological approach to the estimation of returns to research activity. It does not, however, involve those complex problems of evaluating benefits which generally confront the researcher in the health care sector. In addition, one could argue that the research costs for hybrid corn were more easily estimated than would generally

1. T. W. Schultz, The Economic Organization of Agriculture (New York: McGraw-Hill Book Co., 1953), pp. 114-22.

2. B. R. H. Ewell, "Role of Research in Economic Growth," Chemical and Engineering News, Vol. 33 (1955), pp. 298-304.

be the case not only in the health area but also in agriculture.

B. A. Weisbrod

Probably the second most important study completed in the area of "research on research" is that by Weisbrod in his study of the costs and benefits of the development of a vaccine for poliomyelitis.¹ The following formula is used to calculate the internal rate of returns of the polio-directed research:

$$\sum_{t=0}^T \frac{R_t - [B_t(N_t - W_t) - V_t]}{(1 + r)^t} = 0$$

where r is the sought after internal rate of return that will cause the present value of the benefit stream to be equated to the present value of the cost streams,

R is the research costs,

B is the benefits per case prevented as estimated by the loss occurring if polio is in fact contracted,

N is the number of polio cases should no vaccine have been found,

W is the number of polio cases once the vaccine programme is successfully introduced,

V is the cost of applying the research findings,
 t is any one year, and

T is the terminal year on the horizon beyond which the values of costs and benefits are assumed negligible.

Weisbrod then proceeds to estimate each one of these variables in the face of conceptual and data problems. These are particularly severe in the

1. B. A. Weisbrod, "Costs and Benefits of Medical Research: A Case Study," Journal of Political Economy, Vol. 49 (May-June, 1971), pp. 527-44.

about the costs of the vaccination process itself are made.

In comparison with the Griliches work, the work by Weisbrod runs into the foremost problem confronting analysts in the health care field, namely, "How can the benefits be reasonably assessed?". Weisbrod is not able to draw on market prices to evaluate the benefits of the vaccine as Griliches was able to do with the corn. Though the methodology for estimating the benefits accruing from a reduction of morbidity and mortality is being developed, a standard approach is still not commonly accepted.

On the cost side, the estimation of research expenditures is probably at the same time easier and yet more difficult. On the one hand, research for certain classes of disease and the corresponding body systems has been segregated either because the bulk of the funding comes through a particular research fund set up for the specific purpose or because in the adjudication of research grant applications it has been common practice to segregate applications into broad areas to facilitate their evaluation through the peer review system. On the other hand, however, the results of non-mission oriented, basic research and the spillovers from mission-oriented research carried on in widely disparate sectors of economic activity are frequently seen to have significant impacts on the advancement of medical science and technology. Thus, the portion of expenditures on basic research in the economy at large as well as the portion of other research support that should be included in weighing the net value of any given breakthrough in the health research sector will likely have to be arbitrary.

T.R.A.C.E.S.

In 1968, the Illinois Institute of Technology Research Institute published its Technology in Retrospect and Critical Events in Science

(T.R.A.C.E.S.)¹ and in so doing provided a major contribution to our knowledge of the processes by which research efforts are transformed into value research output. Reported on in this study are what would seem on a priori grounds to be the major characteristics of the processes involved in the final outcome of what have proved to be five major scientific breakthroughs. These were magnetic ferrites, the video tape recorder, the oral contraceptive pill, the electron microscope, and matrix isolation.

Though only one of these is wholly an outcome of relevance to the health care industry (the others are said likely to have varying degrees of significance to the delivery of health care), substantial conclusions could be drawn because of the similarity of the processes through which the ultimate outcomes developed. For example, the often hypothesized major role of basic science seemed to be given generally strong support; "of the key events documented, approximately 70 percent were nonmission research, 20 percent mission oriented research, and 10 percent development and application."²

Furthermore, the activity involved in the basic research had two distinct time frames and were likely characterized by two distinct motivation thrusts. In the total period of some fifty to sixty years of research leading to each outcome there was in the first instance a period of undirected, curiosity-motivated, non-mission oriented, basic research. This stage covered some forty to fifty years. There was then a second stage of mission-oriented, applied and developmental research, amounting to some ten years;

1. Illinois Institute of Technology Research Institute, Technology in Retrospect and Critical Events in Science, Vol. 1 (IIT Research Institute, 1968).

2. Ibid., p. iv.

This importantly also involved additional basic research effort to fill in various gaps in the knowledge base of the applied researcher. Indeed, once the mission-oriented research had begun there were still to be accomplished. This suggests that of all the major events forthcoming during the final ten year period of mission oriented research, some twenty percent stemmed from basic science research being carried out in that period.

Though the analyst should with cause be weary in dealing with "numbers of events" and so on, the T.R.A.C.E.S. study does represent a major piece of work. At the same time it serves as much as any study done thus far to underline the infancy of research work in this area.

Interactions

As a follow-up to the T.R.A.C.E.S. study, the Battelle Columbus Laboratories completed a more thorough study¹ (hereafter referred to as Interactions) of the decisive events in the invention-innovation processes involved in the ultimate introduction into the economy of ten innovations. These were the heart pacemaker, hybrid corn, hybrid small grains, Green Revolution wheat, electrophotography, input-output economic analysis, organophosphorus insecticides, oral contraceptives, magnetic ferrites, and the video tape recorder, the last three of which were also included in T.R.A.C.E.S.

Of the several factors evaluated for their influence in bringing about a decisive event, two were seen to stand out, namely "recognition of technical opportunity" and "recognition of need". Ranking first among the twenty-one factors studied, the former was said to be highly or moderately important in 87 percent of the decisive events. "Recognition of need" on the other hand was

1. Battelle Columbus Laboratories, Interactions of Science and Technology in the Innovative Process: Some Case Studies Columbus: Battelle Columbus Laboratories, March 1973).

said to highly or moderately important in 69 percent of the decisive events. The "market pull" hypothesis about the direction of innovative activity thus receives support. It should be emphasized that these factors and others were examined with respect to the eighty-nine individual decisive events rather than to the ten innovations.

With respect to the overall innovation process, the "technical entrepreneur", one who champions the potentially innovative activity was an important characteristic of nine improvements. Early recognition of need was similarly characteristic of nine innovations. Availability of funding, especially of government funding was also judged important as was the occurrence of an unplanned confluence of technology. Finally, it is interesting to note that most of the innovations originated outside the institution that ultimately developed them and that once the innovation process was well launched within the innovating institution further supporting inventions were required in every case.¹

III. RESEARCH ON THE NATURE OF RESEARCH PROCESSES

Our understanding of the processes by which significant breakthroughs in medical science are achieved is dismally low. In the following, we raise some of what appear to be the major gaps in our knowledge base. The majority of these would seem to require a number of in depth case studies of the research processes from which different major breakthroughs have come and of the research activities of various research individuals, groups and centres. These studies, because of their scope are not expected to yield results that would be of immediate use in the formation of government policy in the health research sector.

Alternative Criteria for Making Allocative Decisions

At least three rules have been advanced for making allocation decisions

1. This research method is similar to that used by K. Deutsch, et. al. in their study of 62 major advances in the social sciences over the period from 1900-1965. See K.W. Deutsch, J. Platt and D. Senghaas, "Conditions Favoring Major Advances in Social Sciences", Science, Vol. 171 (February 5, 1971) pp. 450-59.

for research. The first of these states: "Let the researcher be funded to do whatever interests him"; the second: "Let knowledge be found that we do not now possess." The defense of these two criteria is best stated by M. Polanyi as follows:

"So long as each scientist keeps making the best contribution of which he is capable, and on which no one can improve, we may affirm that the pursuit of science by independent self-coordinated initiatives assure the most efficient possible organization of scientific progress. And we may add, again, that any authority which would undertake to direct the work of the scientist centrally would bring the progress of science virtually to a standstill... I appreciate the generous sentiments which actuate the aspiration of guiding the progress of science into socially beneficial channels, but I hold its aim to be impossible and nonsensical... Any attempt at guiding scientific research towards a purpose other than its own is an attempt to deflect it from the advancement of science."¹

In spite of the ringing advocacy of these criteria, they would seem to offer potentially severe limitations if used in allocating government monies to research. Should the researcher's ranking of the gaps in knowledge and/or his own interests in research be markedly different from those commonly held in society resource misallocation would be likely. On the other hand, the individual researcher might be the one best able, if not the only one able, to estimate the probability of his success in any given research area. If this is so, an allocation system that pays little heed to his interests may also lead to resource misallocation.

The third rule advanced² is that the basic economic criterion be used, that the expected benefits of research however roughly estimated should be larger than the costs of research. Though this third rule would

1. Michael Polanyi, "The Republic of Science", in Criteria for Scientific Development, edited by E. Shils (Cambridge: M.I.T. Press, 1968), 6 and 9.

2. For example, see Fraser, op. cit.

seem the only sensible one, it must of course be qualified by another application of the basic economic criterion, namely that the costs of gathering a marginal unit of information should always be less than the expected benefits in order to justify the gathering of that last unit. Thus, it could well be that the costs of assembling the information required to even roughly assess the costs and benefits of research are too high relative to following some other simple rough rule of thumb in allocating resources to the research sector.

In order to assess the usefulness of these alternative criteria it seems necessary that much more research be done of a kind that traces the sources of major breakthrough. The work done as reported in Interactions and as briefly described above needs repeating with respect to other breakthroughs in the health field itself. The work of Jewkes, Sawers, and Stillerman as reported in their book, The Sources of Invention,¹ would seem to provide another important example of the work that needs to be done in order to adequately assess the claims made by scientists like Polanyi as quoted above.

Diminishing Returns to Activity

With respect to research in most aspects of the health care sector, it seems likely that there would be an initial stage of increasing marginal product as additional personnel are committed to research activity; after some level of research involvement, however, marginal product is likely to fall. Over the initial range, the increasing returns corresponding to increments in the volume of resources can perhaps be associated with the division and specialization of research resources, both human and physical resources,

1. J. Jewkes, D. Sawers, and R. Stillerman, The Sources of Invention, (London: Macmillan and Co, 1968).

and with the benefits resulting from the interaction of researchers. However, after some such initial period, it is likely that additional increments in the resource used in research activity will yield diminishing returns. The onset of this diminution is probably directly related to the proportion of the population that is capable of productive research activity. As persons who are not so capable are added to the overall research effort, the corresponding increase in the value of the research output becomes smaller and smaller.

Just which level of research activity exists in Canada at the present time is difficult to determine. However, the above concept of diminishing returns may be useful in explaining why with relatively small resources scientists in the United Kingdom are often said to produce many valuable results while many researchers with vast resources in the United States appear to produce relatively smaller amounts of valuable research results. Though difficult to carry out, a thorough analysis of the health research effort in Canada to determine the nature of the underlying processes of research output would be of some value in view of the tendency on the part of major federal granting bodies to distribute the support rather widely. In this regard we might also recall the results of B.A. Weisbrod, reported above, namely that the return to the last dollars spent on the applied and developmental research seemed to yield spectacular rates of return. But is this the case more generally?

Big Science versus Little Science

Intuition suggests that the gains from concentrating resources are likely to be greater as the research activity becomes more applied, developmental or mission oriented. Conversely, with basic research, the gains

are likely to be largest when research activity is widely diffused even though such diffusion may involve large amounts of duplication.

In this regard we might consider the following assessment: "Knowledge grows through the free sweeping curiosity of the scientist, which if controlled from above will be destroyed. This is particularly true of medical research. For in medicine the problem always seems to present itself in a clear-cut and urgent form: people are suffering or dying from identifiable causes. Hence the frequent earmarking of enormous funds for a direct assault on this or that disease. But to set a beeline for the target may not be the likeliest route by which to reach it. If past experience is to be the guide, public support for medical research should be widely scattered and preferably administered through semi-autonomous scientific bodies likely to take a broad view of the worthwhile fields of enquiry."¹

Clearly there exists an important area for research. What are the likely gains from the concentration of research resources in particular centres? Are the spillovers from one scientist's work to that of another generally significant or just in the more mission-oriented type of research?

Relative Size of the Health Research Sub-Sector and the Proximity of the United States

The near perfect mobility of health research results produced in the United States to the Canadian environment, if this is indeed the case, seems to suggest that the proportion of resources allocated to health research in Canada should be less than that found in the United States if the economic decision rule stated above is operative. On the other hand, the same relative

1. J. & S. Jewkes, Value for Money in Medicine (Oxford: Basil Blackwell, 1961), p. 45.

level or an even greater level might be argued on the basis of this being an area in which Canadians could concentrate their research resources, because of inherent advantages possessed by Canadian researchers. More likely, the interactions between research activities and those involved both in the provision of health care, especially those more complex and specialized services provided in the hospital setting, and also in the education of future medical manpower might justify a level of research activity comparable to that found in other countries. The pros and cons of adopting the "free-rider" approach to health research cannot be satisfactorily assessed however until these various interactions are more carefully studied.

Relative Emphasis on Basic versus Applied Research

The near perfect flow to Canada of research results generated in the United States, again to the extent this exists, seems to suggest that, in following the economic decision rule noted above, relatively greater emphasis should be placed on applied research in Canada. This follows from a belief that the results of basic research will generally have universal application whereas those of applied research may be applicable to conditions characterizing a specific society only.

The degree of emphasis on applied research must be limited however by the following consideration of the nature of the process whereby applied research yields usable results. Applied and developmental research often requires falling back on basic research to fill needed gaps in knowledge. Thus, ten percent of the non-mission research was still to be done when the applied research began in the projects studied by I.I.T. Research Institute, as noted above.

In addition, some minimum number of highly trained scientists may be required to recognize and transmit the useful results of basic research carried out in other countries.

Before sound policies can be formulated on the nature of the balance between the funding of applied and developmental research, a much firmer knowledge base is required on the nature of the interactions between basic and applied researchers.

IV. RESEARCH FOR MORE IMMEDIATE GOVERNMENT POLICY FORMATION

Decisions on the overall allocation of resources to the health research sector and on the allocation of resources within it must of course be made continually on the basis of present information on the nature of research processes. The knowledge gaps discussed above and the research required to fill them will only have a long term impact on policy. There are, however, other gaps in our knowledge base which could likely be filled in the near future. These several areas are now discussed.

The Allocation of Resources to the Health Sector

Of the two fundamental allocative decisions that must be made with respect to the health research sector, the one sets the overall size of the sector. Ideally, one would like to apply directly the cost-benefit criterion advanced above. However, research is by its very nature uncertain and thus precludes the detailed application of this criterion. It has recently been argued by Fraser,¹ in the context that information desired for making a rational decision should only be sought up to the point at which

1. Fraser, op. cit., see especially Chapter VI.

the expected benefits of an additional amount of information are equal to the costs of obtaining this additional amount of information, that decisions on the size of the health research sector should involve the use of guidelines drawn from several sources of information. For Canada and Ontario, the following information was argued to be relevant:

1. The size of the potential benefits derivable from successful health research as seen in either or both of reductions in the direct costs of providing health care and reductions in the indirect costs of disability, morbidity and premature death is extremely large;
2. The comprehensiveness of the definition of health research used as one which includes research activity essential to the operation of an industry (be it in the public or private economy) but which is customarily excluded from standard definitions of R + D in the private economy, such as marketing research, suggests comparatively high levels of funding;
3. the comprehensiveness of the definition of the costs of health research used as one which includes the indirect costs as well as the direct costs of research also points in the direction of comparatively high funding levels;
4. the relative quality of health scientists relative to non-health scientists as judged by non-health scientists seems to suggest the possibility of higher success rates in health research other things being equal;
5. the level of research activity in other industries characterized by rapid technological progress in Canada and in other countries for which comprehensive data are available suggests a target substantially greater than the present experience, namely one in the order of four to five percent of all expenditures on health care;

¹ It is important to consider both of these components of potential benefits since much research may well lead to increases in direct costs yet still be of significant value to society.

6. the relative size of the health sectors in other countries for which comprehensive data are available also suggests a substantially higher allocation to health research, namely one in the order of 4 to 4.5 percent of expenditures on all health care; and
7. the combination of the relative size of health research as compared to the total research effort of several countries and the goal of three percent for expenditures out of GNP on research found in several well-developed countries also points to a target size of about four to five percent of all expenditures on health care. Since health research commonly accounts for up to ten percent of overall research expenditures but at the same time expenditures on health care amount to some six percent of GNP, a target size of 4.5 percent of total health care expenditures devoted to health research is implied if a goal of three percent of GNP is selected for expenditures on all research; and indeed this latter goal has been adopted widely.

Since these guidelines drawn from several sources seem to point to a target of four to five percent for expenditures on health research, Fraser concluded that subject to further research, which was itself strongly recommended, for the funding of health research, including direct and indirect costs, a target equivalent to four to five percent of the total expenditures on health care should be set for the present.¹

Now if one grants that the very nature of research does require the use of guidelines drawn from diverse sources of information, it is essential that the more important of these sources be better developed. In particular, greater effort might with advantage be directed to gathering an accurate descriptive body of data not only on the size of the health

1. An implicit assumption is of course that such funding would be subject to the availability of qualified research personnel and well designed research projects.

research sector in Canada, but also on those of other major, well-developed countries.

Use of Available Technology versus the Production of New Technology

The decision to make more widespread use of available technology and producing new technology must of course be part of the decision determining the overall size of the health research sector.

Now it does seem possible that within a relatively short time span, information on the extent of unemployed but employable advances in medical science and technology could be assembled and analyzed with a view to aiding the policy maker with this particular decision.

There is an additional aspect of this problem, namely the matter of the decision regarding the introduction of a newly discovered health good or service into the health care delivery system. From the vantage point implicit in the above discussion, it often seems to be the case that there are advances in medical science and technology that could with advantage be introduced into the health care system immediately. That they are not is likely the result of a number of factors not the least of which might well be the nature of the health care delivery system itself with the many independently practicing solo practitioners and thereby the potential barriers to a smooth flow of information on the latest advances. In contrast, it is sometimes argued that some "advances" are introduced into the health care system before it has been adequately demonstrated that they are indeed advances. The annual health check up, the Pap smear test, heart transplants may be examples of this claim. In any case, it seems clear that there is an important gap in our research capability, namely that to deal with the

borderline between research per se, and especially its applied and developmental aspects, and decisions regarding the introduction into the health care system of new goods and services. Furthermore, this is a gap that can likely be plugged satisfactorily only with the assistance of the research efforts of economists.

General Myopia and Support of Basic Research

It is often argued that the population at large is short sighted and thus that the politician might well be led to favouring projects with a short term pay-off, that is, those that are very much applied, mission-oriented projects. It is then further argued that scientists such as those on various health research committees must guard against any undue imbalance in favour of short run, applied projects.

In spite of the intuitive appeal of this argument, so often advanced by scientists, an attempt to assess the available evidence would seem in order. Thus, the actual revealed preferences of various funding bodies and their peer review panels might be examined with the goal of attempting to evaluate the nature of the influence brought to bear on the funding decisions of scientists, government representatives, politicians and any other interested parties.

Grants in Aid of Research versus a System of Prizes

One of the fundamental questions that must be asked is whether the system of funding health research that presently characterizes much of the health research sector in Canada is the appropriate one. It is principally a system of grants-in-aid of research. Such grants are in the main given to individual researchers for specific projects; there is also some "programme"

funding for groups of individuals working on a more or less coordinated set of research projects.

An alternative system that might be considered would involve a reorientation of the grants-in-aid of research from individuals and their projects to groups and their programmes. In particular, one might well consider the selective use of health research institutes in which the vast majority of the health research would be done, as is the case in Czechoslovakia.

Another system that could involve the complete renovation of the existing one would be one based on a network of prizes for work already accomplished, wherein the value of the prize would be determined by the value to society of the outcome. Though the difficulties involved in evaluating outcomes, even in those cases in which the outcome is seen to have merit within the lifetime of the researcher in question, are legend, such a system is worth considering. This seems especially so since the grants in aid of research systems often involve the use of committees which are frequently alleged to be the captives of present day fads and misconceptions.

V. A RESEARCH AGENDA

The present state of the knowledge base is such as to all but preclude the setting of anything but an arbitrary set of priorities for research in this area. We can start with the basic premise that health research is a good thing not only because of the benefits received today, principally in the form of spillovers from research activity to the delivery of health care and the education of medical manpower, but also because of the potential benefits likely to accrue in future years to the present population and future generations of it. We are then faced with the fundamental economic question:

"What and how much of the many potential kinds of health research activity should be carried on in Canada?". An answer to this question is continually being given through the policies of government and the actions of individual researchers and other involved parties.

The quality of the answer could, however, be improved by research effort directed to some of the major policy issues and knowledge gaps that we have discussed above. There are four areas in which an immediate research thrust could be made with probable success. These are first, the assembling of an adequate data base on the size and structure of the health research sector in Canada and in several major economically well-developed countries;¹ second, the start on a programme of case studies of the research processes underlying a number of major breakthroughs in medical science and technology; third, the start of a programme of case studies of the research activities of what are seen by their peers to be successful individuals, groups, and institutes in the health research field; and fourth, the start of a comprehensive series of studies of particular health goods and services that have recently been introduced into the health care delivery system with a view to determining first whether their introduction has indeed proved to be of net value to society and secondly, whether their further development might prove advantageous.

1. A start on this has been made by Fraser, op. cit., see especially Chapter II.

CHAPTER IX

THE DISTRIBUTIONAL IMPACT OF GOVERNMENT FUNDED HEALTH CARE

I. INTRODUCTION

Income Distribution as a Program Objective

Universal health insurance schemes are frequently justified on the grounds that good health is a universal human right and that to realize this right some redistribution of real income must occur since some individuals have insufficient income to sustain themselves and their families at a reasonable level of health. In addition it is usually assumed that only by distributing health care directly will the financial barriers effectively be removed, as opposed to use of the theoretically more efficient method, of distributing dollars to the needy.¹

But this general rationale for redistribution is insufficiently precise and has led to considerable confusion among policy makers and health care analysts alike. As a result the goals of health service programs have been poorly specified and understood.

Three quite different ideas govern or help in setting policies for various mechanisms of health care provision: the income tax mentality, the minimum income mentality and the social fault mentality. The first of these has to do with ability to pay criteria and equitability, the

1. These and other basic theoretical problems are introduced and discussed by (a) Kenneth Arrow, "Uncertainty and the Welfare Economics of Medical Care", *American Economic Review*, Vol. 53 (December, 1963), pp. 941-73; and (b) Mark V. Pauly, "Medical Care at Public Expense: A Study in Applied Welfare Economics", Ph.D. Thesis, University of Virginia 1968, pp. 7-19.

* The research assistance of Mr. T. Edwards, an Honours B.A. (Economics) Graduate of Queen's University, in assembling the background material for this Chapter and in setting out the first draft of it, is gratefully acknowledged.

second is concerned with assuring a certain minimum level of care for all while the third recognizes social responsibility for certain segments of the population. Although these approaches are often used interchangeably, they have divergent implications for the degree and methods of achieving a redistributive impact. For instance, the income tax mentality only operates with respect to income groupings according to their ability to absorb health care expenses while the equal access objective not only removes income barriers for the needy but also raises consumption barriers to the wealthy and the social fault mentality identifies socially and geographically isolated groups to whom special consideration is given.¹

Canadian legislators have made the above statements of objectives as well as more glamorous if less attainable ones. The Royal Commission Report on Health Services of 1965 went so far as to claim that "as a nation we (should) now take the necessary legislative organizational and financial decisions to make all the fruits of the health sciences available to all our residents without hindrance of any kind".² This profession of good intentions pays little heed to practical constraints and is quite contradictory to a following statement, taken from the same report, that defined universal health care as meaning that "adequate health services shall be available to all Canadians wherever they reside and whatever their financial means".³

1. Robert J. Lampman, "Comments on the Political and Economic Limits of Health Insurance", *Bulletin of the New York Academy of Medicine*, Vol. 45, No. 8, (August, 1969). See also Irwin Garfinkel, "Equal Access, Minimum Care", *Journal of Human Resources*, Vol. 7 (1972), pp. 242-249.

2. Royal Commission on Health Services, Queen's Printer, Ottawa 1965, Vol. 1, p. 10.

Ibid., p. 11.

Of course, the actions of consumers and providers of health care as well as government policies, or lack of them, have determined the distributional consequences of the national health scheme; but it is probably safe to say that neither private individuals nor the government even yet know what the impact has been. This ignorance is perhaps the result of a lack of investigatory effort, and, or, a lack of data. In addition, there is little chance, beyond the sheerest of coincidence, that the distributional impact of the national health services program has, in fact, met with the vague and ambiguous intentions of the Hall Commission.

These ill-conceived program objectives have been magnified at the provincial level where inter-provincial inequities are extreme and intra-provincial correlation of objectives and consequences remain uninvestigated.¹

Definitions and Limitations

Having noted the range of distributional objectives the question arises as to how one might establish the ultimate income distributional effects of each. In framing this discussion we should define the area under consideration. First the programs included in this kind of analysis are normally those contributing directly to the physical and mental health of the population (i.e., sanitation, education and recreation await a more ambitious appraisal). Secondly the phrase "income distribution" refers to the actual dollar or in kind benefits and payments received and made by individuals; the emphasis being on a specification of who pays for and who receives.

1. N. Stefanuk, "Comparison of Provincial Hospital Plans", *Canadian Hospital*, Vol. 46 (1972), pp. 82-88.

substitution of more-expensive for less-expensive manpower; e.g., obstetricians for midwives, the continuing support of an expensive licensure system and the use of a reimbursement system that offers few economic incentives for efficient operation. Accordingly the growth of auxiliary personnel has not been sufficient to reveal a pattern of substitution comparable to that of other industrialized countries. For instance, the ratio of dental hygienists to licenced dentists in Ontario was 1 : 9 in 1967 and 1 : 7 in 1972. Interestingly, the growth rate of dentists exceeds that of hygienists, even though each dentist could provide work for at least one hygienist.¹ So, too, for physicians, the rate of growth in their numbers over the last five years in Ontario has averaged 7.9 per cent. This is greater than the growth rate of nurses and orderlies and almost as large as that of nursing assistants. Not only have the usual types of manpower not been substituted but neither have new types. Beyond a few demonstration projects, physicians assistants and pediatric nurses are almost unknown. In addition, it should be noted that the distribution of health personnel in Canada is by no means uniform with respect to population. From a regional point of view, the range of dentists per capita in 1972 was from 1 : 8, 313 in Newfoundland to 1 : 1,076 in Prince Edward Island to 1 : 586 in Ontario.² These rough figures suggest that Ontario and Canada residents are paying excessively for health care by supporting inefficient production methods and that the care received could be better distributed.

1. Profiles on Health Occupations, Ontario Department of Health Research and Planning Branch, 1969 and Canadian Manpower Inventory 1973, Dept. of Health & Welfare, Canada.

2. Canada Manpower Inventory 1973, Dept. of Health & Welfare, Canada.

In the following section we will be chiefly concerned with the possibilities for improving the efficiency of production methods. The organization of this section will be to first identify the production function for health care, and to discuss the nature of health care input and output. Secondly, we wish to discuss potential areas for productivity growth and the constraints to such growth. Thirdly, we look into some of the major policy issues to which economic analysis can be readily applied. Fourthly, some selected case studies of substitution possibilities will be reviewed with the intention of demonstrating the potential for substitution. Finally some suggestions regarding research proposals will be made.

II Production Function Approach

In this section we review the possibilities for manpower substitution in the production of health services. From an economist's point of view this involves an analysis of the health sector production function, focusing on manpower productivity. A generalized production function for the health sector might be written:

$$Q = f(L_1 \dots L_n, K_j \dots K_n, T)$$

where Q is the quantity of health services produced by various combinations of L_i , health manpower inputs (physicians, nurses, technicians), and K_j , capital inputs (buildings, equipment, disposables) at a given level T , of

them to less appealing localities or that find substitutes for the current sources of primary care; for example, nurse practitioners.

The final major area of policy involves the regulation of the consumption of health care. The mechanism used today falls heavily on the shoulders of health care professionals who neither have the knowledge of how important this responsibility is or of how to carry it out in a socially optimal fashion nor the incentives to do so. In fact, as suggested in the demand-supply model presented later on, the incentives that health professionals currently face lead to outcomes contrary to program objectives. Until we are in the enviable economic circumstance of being able to provide everyone with perfect health at near zero costs there will be a need for methods of restricting demand. In any case, there now exists a very pressing need for control mechanisms.

Methods of controlling health care consumption can be grouped under three headings: copayment schemes, supply restrictions, and consumer based controls. Cost sharing or copayment takes four major forms. Under a deductible scheme the consumer pays all expenses up to a certain point, where the insurance coverage begins. With a coinsurance scheme, the consumer pays a specified percentage of the insured expenses. Under a deterrent fee mechanism a fixed fee is charged per unit of insured service consumed. Finally indemnity payments provide a fixed payment to the consumer per unit of service used regardless of the actual charges.¹ As suggested by M. Feldstein these mechanisms can all be reviewed with

1. Paul B. Ginsburg and Larry M. Manheim, "Insurance Copayment, and Health Services Utilization: A Critical Review", *Journal of Economics and Business*, (Fall, 1973), pp. 142-153.

regard to their impact on the utilization rates of various socio-economic groups in examining income distribution aspects of each method.¹

Supply restrictions have also been used to limit consumption.² The idea is that simply by limiting the number of doctors, dentists and hospital beds there will be a very real limit placed on the amount of services consumed. Unfortunately the income distributional impact of this method is not likely to be very attractive. A study of Montreal's health care consumers revealed that the reduced availability of physicians' services as measured by longer waiting periods for appointments and in offices led to a redistribution of services from the aged and the very young to middle age groups. The reason suggested for this was that the very young and old are less able to stake their claim on scarce resources. In addition to supply restrictions of this variety it is conceivable that adjusted income incentives for health care professionals and educational programs for these professionals could improve the current system a good deal. One further supply restriction has to do with the services that are insured. It is reasonable to suppose that income distributional goals would be served more directly by an analysis of the services that needy groups consume. We have seen little in the way of research along these lines but the concept is repeatedly implied by figures showing high incidence of venereal disease, mental illness and tuberculosis among the indigent.

1. Martin Feldstein, et.al., "Distributional Aspects of National Health Insurance Benefits and Finance", National Tax Journal, Vol. XXV, #4 (December 1972), pp. 497-510.

2. For a discussion of consumption controls of the nonprice variety see K.J. Arrow, "The Economics of Moral Hazard: Further Comment", American Economic Review, Vol. 58 (June, 1968), pp. 537-39.

Another consumer based demand control is the development of a system for dealing with hopeless cases. At the present time a very high percentage of hospital beds are occupied by patients who for all intents and purposes are merely existing. This is a very touchy research area but one that must be approached. Also consumer education with the intention of reducing consumption might be effective. As Marc Lalonde has noted, the fact that health insurance reduces the price of care to zero may have taken a good deal of the responsibility from the consumer for his own health. Somehow this responsibility must be returned.

The general implications for all these methods of consumption control are that they are likely to be nominated as an important method of curtailing rapidly rising health care costs, and they will each have very important income distributive impacts. Accordingly the methods chosen will have to be very precisely investigated for their overall impact on utilization rates and income distributional consequences. The major policy decision being which ones of these controls to use and to what degrees.

In the following section on the income distributional aspects of the health sector we will, in the light of distributional objectives described above: a) discuss the nature of the distributive impact; b) point out some major policy decisions that dictate distributive consequences; c) review the relevant studies carried out in Canada on various distributional aspects and d) develop an agenda of research projects.

II. THE NATURE OF THE DISTRIBUTIVE IMPACT

That real income is redistributed as a result of health service programs is partially due to the successful achievement of program objectives and partially due to side effects of adopting a particular structure for organizing the health care delivery system. In the discussion below we will identify some of the major mechanisms through which redistribution occurs. Our main concern, however, is with the more direct redistributive consequences.

Cost Burden of Health Care

The cost side of health services is distributed among community members by health insurance premiums, by all manner of taxes and by government debt financing. The premiums account for about one third of the total health insurance scheme and the general tax revenues for the remaining two thirds. An analysis of the incidence of these general tax revenues involves two principal considerations. First it must be recognized that in the end the entire tax burden must be borne by individuals. Though taxes may be collected from business firms their ultimate burden must be traced to individuals or households, in their capacity as owners, as employees or as consumers of products. Second, the final burden distribution may differ from that of statutory liabilities whether the tax is imposed on individuals or on firms. Individuals as well as firms may adjust their sales and purchases, thus affecting the position of others. Determining the actual distribution of the tax burden thus requires an analysis of the economic adjustment process, or the transmission of the burden from its impact point (the place of statutory incidence) to its

final resting point (the place of economic incidence). This process is generally referred to as "shifting". It is obviously the distribution of the taxation burden after shifting that is at issue. This task has been started by W.I. Gillespie and I.J. Goffman for Canada but there is a pressing need for a specific study of the incidence of the tax burden in Ontario. In the absence of this study it is generally accepted that the combined federal provincial and local tax burden is slightly progressive.¹ Specific to medical expenses, the medical deductability for income tax purposes and the standard \$100 medical deduction introduces a good deal of progressivity into the health sector.²

As for the premium incidence of health insurance schemes this is a straightforward design characteristic that can be manipulated according to the objectives of the program. For instance in Ontario anyone over 65 or earning less than \$1,000 in taxable income may apply for full premium assistance while partial premium assistance is available to individuals of slightly greater earnings. The major distributive impact here is the fact that all individuals not receiving premium assistance must pay for the health services consumed by the assisted population either in the form of increased premiums or increased taxes. One devious problem with this method is that of "notching" where a lower income group

1. Irving J. Goffman, "The Burden of Canadian Taxation: Allocation of Federal, Provincial & Local Taxes Among Income Classes", *Canadian Tax Foundation* (July, 1962). See also W.I. Gillespie, "The Incidence of Taxes and Public Expenditure", *Royal Commission on Taxation*, Vols. 2 and 25.

2. It is not clear why this deduction has been maintained given the coverage for health expenses that now exists through insurance.

is leapfrogged over the marginally higher income group that has to pay the premiums. But a more important problem with the premium system of the Ontario insurance plan is that there is no correlation between the premium level and the risk of ill health. The rationale behind any insurance is, that by spreading the risk of an unpredictable and costly life experience over a large group of people, all will benefit. But this is only true in so far as groups can be identified as having equal likelihood or experiencing the risky event and that an actuarially fair premium can be calculated. If one group of individuals has a greater probability of suffering ill health than another group, and if they pay equal premiums, there will be a redistribution from the non-users to the users.¹ In Ontario this arrangement is taken to the extreme that all individuals are insured under equal premiums and are, therefore, assumed to be equally likely to experience ill health. If we are able to determine who, in terms of socio-economic, demographic and regional characteristics, is, in fact, likely to experience ill health and therefore likely to use the health services made available, it would be possible to calculate the degree to which risk spreading is redistributive.

Beneficiaries of Health Care: Ex-Ante Versus Ex-Post Health Status

There are both direct and indirect beneficiaries of health service programs. The latter are dealt with separately in a later part. The incidence of the former can be viewed in terms of the health status of individuals prior to contacting the health sector, the availability of

1. Kenneth Arrow, "Uncertainty & Welfare Economics..." op. cit.

health care resources to them, their utilization of the resources available, the quality of the services then rendered, and finally in terms of the final health status of the individuals in question. Persons of different socio-economic, regional and demographic characteristics are expected to receive different amounts of benefit from the health care sector for these several reasons. With objective measures of health status before and after entry to the health care delivery system it would be possible to set out an adequate description of the beneficiaries of government funded health. Such objective measures are however not available both because of conceptual difficulties and because of limitations in the availability of data. Our description of the beneficiaries of the health care system thus involves a consideration of the intermediate stages in the process of the delivery of care. Before beginning this consideration we might first examine the problems of setting out a measure of the "value added" to an individual's stock of satisfactions by the health care delivery system.

To determine the condition of the raw human material prior to the health care process and to evaluate the quality of the finished product following the process, we would need to be able to develop an index of health status that can be used to identify an individual's "need" for health care. In a recent literature review of health status indicators it was concluded that no workable system was yet available that could accurately monitor the level of health of a large population.¹ On the

1. For discussion see a) S.B. Goldsmith, "A Revaluation of Health Status Indicators", *Health Service Reports*, (December 1973), pp. 937-42; and b) Dania F. Sullivan, "Conceptual Problems in Developing an Index of Health", U.S. National Center Health Statistics (1970), No. 2.

other hand, much research is now aimed in this direction, and policy makers should be aware of developments in this critical area. For the time being some, albeit inadequate, measures are available and are generally referred to as mortality and morbidity health indices, mortality referring to infant mortality rates, age-adjusted death rates, life expectancy tables, or crude death rates. Mortality statistics are used with reservation because, though there are almost the only available measures of health status that are reasonably objective, they do indicate only the all or nothing state of health as in life or death. Morbidity, the measure of sickness or illness present in a population, suffers from a lack of specific criteria and data upon which to base an index of health. Some available but inadequate measures of morbidity are the number of days in hospital and the number of hospital admissions. These broad groupings tell very little about the cause of illness or the degree of disabling effect. Also, they ignore any treatment that is carried on outside hospitals and may place undue emphasis on acute rather than chronic conditions. Furthermore, they are influenced by the availability of beds and as yet cannot be at the margin associated with increases in health. At present, the possibilities for measuring the input and output of health services programs in terms of the health status of individuals looks bleak. Perhaps what is required above all else is a decision that picks one of the many methods suggested in the literature as a temporary standard for data collection and analysis until a generally accepted method becomes available.¹

1. Work done by C.W. Torrance and by the Ontario Ministry of Treasury, Economics and Intergovernmental Affairs should be reviewed in choosing a health status index.

1. C.W. Torrance, "A Generalized Cost-Effectiveness Model for the Evaluation of Health Programs", Research Report, Faculty of Business, McMaster University.
2. Health Index for Ontario, Ontario Ministry of Treasury Economics and Intergovernmental Affairs (December 1972).

The Beneficiaries of Health Care: Availability of Resources

Because of the difficulties involved in setting out estimates of net improvements in health status, alternative methods of estimating the distribution of the benefits of the health care system must be devised. The first of these that we consider is the level of availability of health care resources. Analysis of the influence of availability, carried out recently in Manitoba, lead to the conclusion that the distance travelled to medical facilities is an important determinant of consumption rates and that the availability of services in Manitoba is extremely inequitable.

Data describing the distribution of dentists, physicians and other health care facilities are available. Furthermore, they could be related to census tract statistics in an effort to shed light on the nature of the social, economic, and other determinants of availability. And of course the basic data describing the distribution, that is the availability, of health care resources would provide some insight into the likely distribution of the benefits of the health care delivery system.

The Beneficiaries of Health Care: Utilization

The actual availability of health care resources does not give us a comprehensive view of the receipt of services from these resources since the pattern of utilization of these resources differs as between different segments of the population.

The literature on the utilization of medical services by socio-economic and regional groups is far from complete, but is nevertheless

one of the most investigated areas relevant to our discussion in this section. The following is a brief list of some of the suggestions regarding causality of low levels of utilization by groups of low socio-economic status. Under the general heading of cultural beliefs some of the reasons given are first, fatalistic attitudes and feelings of powerlessness among the poor that resist care; second, preference for ethnocentric social organizations that maintain "popular" rather than scientific attitudes toward care; third, less symptom awareness and concern that leads to reduced use of care; fourth, the tendency toward short-term goal orientation and general ignorance of long-term consequences among the poor; fifth the lack of accurate knowledge with respect to health hazards, or availability of service that prevails among low socio-economic groups. Interpersonal influence is cited as another area of concern. For example, more discussion of ailments is carried on the higher the socio-economic group, relationships between the poor and the medical professionals are often described as being uncomfortable, the difficulties of communication between professionals and poor laymen are greater, and the complexity and impersonality of clinics and hospitals cause strong negative reactions in the poor.¹

The role played by finances per se as a variable connecting socio-economic status with utilization of medical services must, of course, be considered. "The data available, while demonstrating that this factor

1. Martin Hyman, "Some Links Between Economic Status and Untreated Illness", *Journal of Social Science and Medicine*, Vol. 4 (1970), pp. 387-399.

cannot be ignored, refute the belief often expressed by laymen that it is the sole link between socio-economic status and utilization behaviour."¹ The forgoing statement made with reference to American studies may not apply to the Canadian situation. That poverty in Canada may be of a very different variety than that existing south of the border cannot be theoretically judged. In the United States the problem of poverty is certainly of greater magnitude with deep cultural roots giving support to the concept of a "hard core" poverty class. In Canada, some authors maintain that poverty is still a deep rooted problem with nearly 20 percent of families in 1969 being classed as low income by the Economic Council of Canada (actually 17.13 percent). This figure is surprisingly high and deserving of greater analysis. Some of the questions that need to be answered are the following: "How low an income will initiate poverty characteristics?" "Are the families in the group constantly poor or are they being changed regularly (i.e. university students, the unemployed and the temporarily disabled)?" "Does low income in Canada cause the other negative problems associated with poverty; i.e. lack of education, lack of nutrition, alienation and fatalism?" and "Would it be possible to treat this poverty group more directly through regional programs of subsidy and comprehensive environmental change?".

In addition to analysis of utilization characteristics of the poor, it is necessary to investigate all socio-economic and regional groups. The result of this kind of study will indicate to what degree individuals are consuming and how the high consumers correspond to the high payment supporters of the services. Studies carried out in Canada

1. Martin Human, "Some Links...", *op. cit.*, p. 390. See also M.J. Lefcowitz, "Poverty and Health: A Re-examination," *Inquiry*, Vol. 10 (March, 1973), pp. 3-13.

to date yield contradictory results. Dr. Enterline's study of Montreal physician utilization revealed significant redistribution of services available. He also observed disparity of physician utilization with members of low income households making 7.8 visits per year as compared with 4.8 by high income household members.¹ Dr. Beck's study of Saskatchewan physician utilization, though not directly comparable in that only post universal services provision observations are reported, and a different definition of utilization is used, suggests significant disparity of physician accessibility with low income household non-users being in far greater proportion as compared to non-users among high income classes.² Beck also pointed out that disparities of utilization are not likely to equalize after even five years of operation of free services.

That these two studies were carried out in very different settings and used different methods may explain their opposite conclusions. Above all they reveal the need for more in-depth study with precise definition of terms such that the test and the results can be validated in subsequent research. For our purposes it is clear that health care is consumed, for reasons associated with different socio economic and demographic groups at very different rates.

1. P. Enterline, "Distribution of Medical Services Before and After 'Free' Medical Care", *New England Journal of Medicine*, Vol. 289, p. 1174.

2. G. Beck, "Economic Class and Access to Health Services", *International Journal of Health Services*, (Winter 1972).

The Beneficiaries of Health Care: Quality of Services Provided

Once we have knowledge not only of the distribution of the health care resources but also of the patterns of utilization of these resources, we are well advanced in our attempt to estimate the distribution of the direct benefits of health care. One significant additional factor that must be considered however is the quality of the services provided, and in particular to enquire if there are causal links between quality of care and any of the several social, economic, demographic, geographic factors that influence both the availability and use of health care resources.

The question of differences in the quality of service is extremely difficult to answer. Some measures suggested in the literature are number of specialists available, number of physicians available and variety of services available. But not only does this overlap to a large degree with the availability aspects mentioned above, it may not be representative of quality levels. Some surveys have been carried out that actually involved observation by a trained group of professionals attempting to evaluate quality of care.¹ This method suffers from cost limitations and professional objections to "peer review". Also, it does not usually come under the heading of measuring quality distribution that is our concern at present.

1. For a discussion of quality of physicians in Ontario see K.F. Clute, The General Practitioner (Toronto: University of Toronto Press, 1963).

Clearly when a third party payment system is adopted there are a new set of incentives present as compared to private provision systems. For instance, if physicians face a fee-for-service system, then practices in pleasing environments; i.e. middle and upper income, urban areas, will receive not only more doctors, but the so-called cream of the crop, while less attractive areas may receive progressively less capable and less satisfied practitioners. True enough, the question is an empirical one, but one that is deserving of attention in determining income distributional consequences.

In addition to the physician incentives, there are a variety of patient incentives that alter the quality of care. For instance, given the much reviewed environmental constraints that limit the ability of poverty groups to function in a middle class society, it is inevitable that health services should be a likely target for discussion. There appear to be a number of possible reasons why the poor do not receive quality care, for example, because they choose to make care discontinuous; or, because they are not well enough educated to understand the implications of comprehensive care; or because they are not capable of interacting with the health care system in the "required" manner. This last point is the most important in that it should be understood that providing a given standard of quality of care to certain groups may be far more expensive than for other groups.

In general, one might be able to standardize the quality of care provided to a large majority of the population through a system of regulation and supervision, if indeed this was thought desirable. On the other hand, as long as the services available to regionally isolated

areas are of inferior quality there will remain a distributive problem. Further, as long as quality of services cannot be made available to certain groups because they choose to not utilize the service in a "proper" fashion, then the real success of redistributive programs with respect to health care will be constrained.¹

Indirect Distributive Effects

It is difficult to generalize about the indirect distributive repercussions of public provision of a health service program. The following paragraphs will identify some of these in as much as they may outweigh the progressive intentions implied in the foregoing sections.

The first area in which a redistributive impact is likely relates to the monopolistic nature of the organizations that exist in the health sector. From a factor input standpoint there has been little control of the prices of capital or labour. The result is that marginal cost pricing is not adhered to. In the case of the hypothesized monopoly held by medical professionals, the redistribution is from consumers to producers on a grand scale. Not only because higher prices are charged but also because health care personnel have not been equitably distributed on a regional basis. In the case of generally inefficient institutions, the impact is far greater as some authors have argued. They have claimed

1. Anslem Strauss, "Medical Organization, Medical Care and Lower Income Groups", *Journal of Science and Medicine*, Vol. 3 (1969), pp. 143-177.

that hospital inefficiency has led to the 13 - 14 percent annual inflation rate of medical care costs and prices.¹ Here we can identify another redistribution from consumers in the economy at large to producers in the health care sector. Considering the size of the health sector significant macroeconomic implications may also have distributive consequences.

A second source of indirect redistribution can be detected in the federal provincial structure of tax collection and health sector administration. Because the technical arrangements differ for each province the amount of federal support varies from 40 percent in Ontario to 90 percent in Newfoundland. This is not an uncommon characteristic of federally funded programmes but for purposes of an income distribution analysis oriented to Ontario the impact may be quite large.²

Thirdly, health programs that are supported by general tax revenues and cannot be said to fall into the category of insured illness will have a distributive impact. Such programs as public health, health research, and chronic care are supported by government funds in recognition of the need for them but not specifically aimed at distributive goals. Buchanan and Tullock have argued that in fact this is a form of insurance in that the products of these programs insure us all in one way or another

1. Ronal Andersen and John T. Hull, "Hospital Utilization and Cost Trends in Canada and the United States", *Health Services Research Reports*, (Fall, 1969), pp. 198-222.

2. E.N. Stefanuk, "Comparison of ..." *op. cit.* For the American cases of this problem see Bruce Stuart "Who Gains From Public Health Programs", *Annals of the American Academy of Science*, (Fall, 1971), pp. 145-149.

(i.e. taxes used to support institutions for the chronically ill insure us all for the cost of chronic illness).

The distributive consequence of this kind of program is that certain groups will be expected to consume the product of these programs to a greater degree than other groups. For instance, the greater propensity of low income groups for mental illness dictates a high proportion of low income individuals in mental hospitals. Similarly the benefits from medical research may be more applicable to disease among high income as opposed to low income groups. The result is that public expenditure in the health sector has a broadly predictable redistributive impact.

III. A MODEL OF THE DISTRIBUTIONAL IMPACT OF GOVERNMENT FUNDED HEALTH CARE

The foregoing discussion of the nature of the impact of government funded health care has highlighted the problems involved in its estimation. We now wish to set forth briefly an hypothesis of the likely effect of government's participation in health care delivery.

We begin by assuming that the demand for health care differs as between high and low income groups where income is a proxy for the amalgum of social and economic factors thought to be important determinants of care. The demand of the high income group is thought to be relatively high and relatively elastic in the relevant range. For example, trips to Bermuda may be seen as relatively good substitutes for a package of visits to the psychiatrist during the winter months. In contrast, the demand of the low income groups is thought to be both relatively low

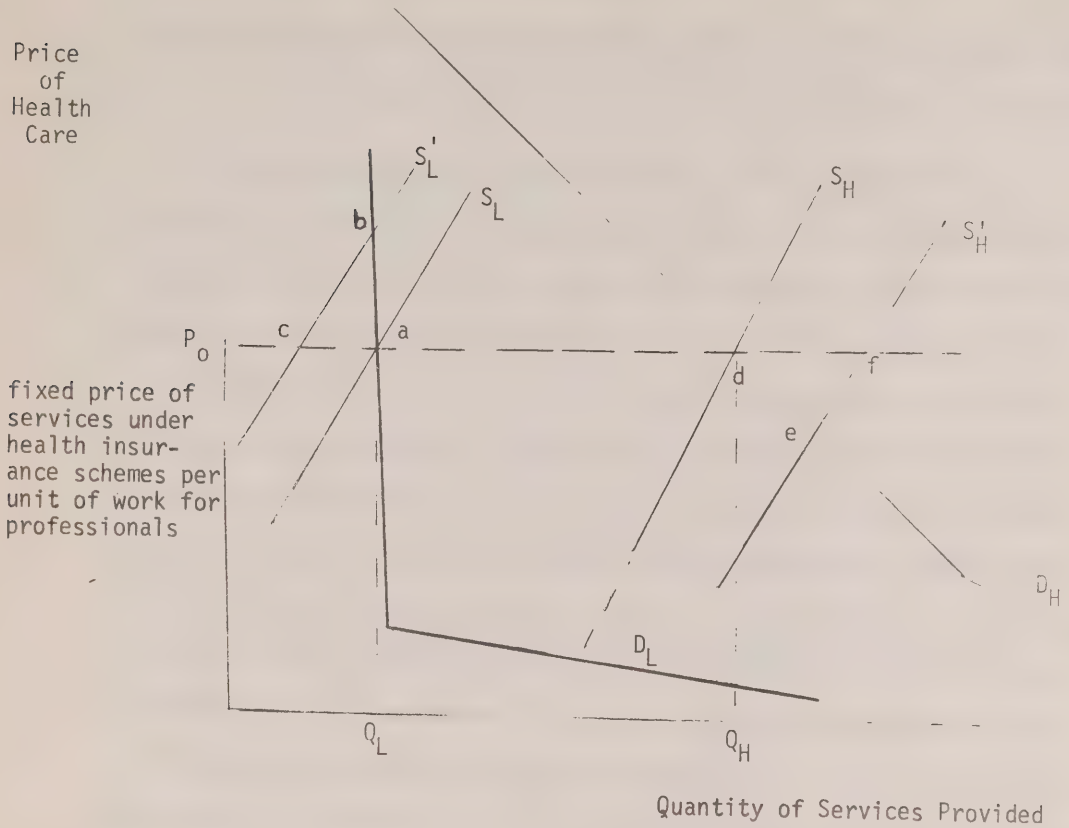
and relatively inelastic. These demand curves are shown by the curves D_H and D_L , respectively, in Diagram IX-1.

On the supply side, in the pre-health care insurance era, the supply of physicians to the low income groups is thought to be relatively lower than that to the high income groups. This difference is thought to be the result of two factors. On the one hand, the psychic income involved in providing services to the high income group is thought to be higher than that to the low income group. In addition, there are thought to be distinct differences in the location of the high and low income groups such that physicians would prefer other things being equal to live in the high income areas rather than the low income areas. Not only will this be satisfactory to their families but also to themselves since the high income areas may also be the preferred sites for the wide range of facilities complementary to the physician such as the hospital, and so on. Supply curves S_L and S_H are drawn in Diagram IX-1 to illustrate these differences prior to the establishment of the health care plan.

Given the use of fee-for-service schedules, equilibrium occurs at the same price in both the low income and high income areas, as shown by the price P_0 in Diagram IX-1. The levels of output differ however from Q_L in the low income area to Q_H in the high income area.

Following the establishment of the government health care plan, it is no longer necessary for the physician to seek out patients under the constraint of their having to back their demands with purchasing power; for these demands are now backed by government purchasing power.

Diagram IX-1 Demand & Supply of Health Sector Personnel



where:

D_L = demand for health care personnel in low income areas

D_H = demand for health care personnel in high income areas

S_L = supply of health services provided by private practitioners in low income districts

S_H = supply of private practitioners in high income districts

S'_L & S'_H = supply curves following introduction of the health care plan

Q_L = quantity consumed by low income families

Q_H = quantity consumed by high income families

Note: income is used here as a proxy for the broad range of characteristics that are thought of as defining poverty groups.

Given that the demand for health care is insatiable, it is possible for any given physician to generate a desired level of income in the location of his choice. We would therefore expect a shift in the supply of physicians away from the low income group which we have associated with poor locations as perceived by physicians to the high income groups, as shown by S_L' and S_H' in Diagram IX-1.

A further consequence is seen on the demand side. Because of the nature of the demand curve of the low income group, government funded health care at a much lower net price¹ than that previously incurred leads to a relatively small increase in the quantity demanded. In contrast, there is a relatively large increase in the quantity of health care demanded by the high income group because of the nature of their demand curves. The net result on the demand side is that the shortages of health services are observed to be greater in the high income areas than in the low income areas.

Since government involvement in the funding of health care has not resulted in increases in the supply of health care resources that would meet even the greater part of the increases in the quantity of health care demanded, the post-related result is one of there being a redistribution of health care resources away from low income groups to high income groups.

This is of course a testable hypothesis. In addition, there

¹The net price is indicated largely by the opportunity cost of the patients time allocated to obtaining health care plus any transportation costs.

are probably sufficient data in existence that if gathered could be used to test this hypothesis. Clearly, if government is interested in assessing the effectiveness of its intervention in the health care field, it is a hypothesis worth testing. Moreover, the model of the supply and demand of health services from which it is drawn is one that could be expanded in a number of ways to account for different aspects of the demand and supply of health services that may have been affected by the introduction of a government funded health care scheme.

IV. LITERATURE SURVEY

- M. Feldstein, et. al.

From a methodological point of view, an article by M. Feldstein, F. Briedman and H. Luft made the first comprehensive analysis of health insurance.¹ Their method simulated the distribution of health expenses according to: one, the structure of the insurance coverage and the way it is related to family income and family composition; and, two, the mix of revenue sources (premiums, income tax, payroll tax) that is used to finance the program. They pay strict attention to dollar flows and in so doing ignore several central issues such as utilization rates, availability factors and quality factors. Benefits are also simulated for the insured group according to demographic factors and the price elasticity of demand (when coinsurance is reviewed). The contribution of this study is seen in the demonstration that the actuarial value of the benefits accruing to each family under different price elasticities and different coverage structures can be calculated. Secondly, these authors identify four measures of the income distributional

1. Martin Feldstein, et. al., "Distributional Aspects...", op.cit.

aspects of each insurance scheme in so far as these measures point out slightly different facets of income distribution. First, they analyze the progressivity of the program by a straight forward cost benefit analysis. Secondly, the implied marginal "net tax" rate is calculated by expressing the difference between costs and benefits as a percentage of income. Then the extent of protection against relatively large medical expenses is reviewed in so far as the program objective may be to insure all income groups against large medical expenses while leaving the responsibility for minor expenses with the consumer. Finally a measure called the "uniformly distributed dollar" is described which combines the benefits to families in different income classes into a single measure of distributional equity. This is a weighted sum of the benefits per family in each income class where B_i is the benefit per

$$UDD = \frac{\sum_{i=1}^N B_i W_i N_i}{\sum_{i=1}^N W_i N_i}$$

family in income class i , W_i is the weight given the marginal dollar to a family in that income class (marginal social utility of a dollar given to a family in that income class) and N_i is the number of families in class i . Thus one UDD unit is the social value of uniformly distributed dollars.

With this data base and for the various insurance schemes examined, they demonstrated a) that the range of income distributional effects is quite large depending on what coverage is used; b) that the program objective will be satisfied in varying degrees as measured by the four characteristics identified earlier; and c) that this kind of income distribution research

as a preliminary analysis will demonstrate the apparent distributive impact but will have to be further adjusted by utilization and availability factors.

P. Enterline

There have been five studies carried out in Canada with respect to the effect of national health insurance on physician care utilization. The most comprehensive of these was carried out in Montreal using a before-and-after questionnaire survey by personal interview of 5789 households.¹ Information obtained included demographic characteristics, illness type and severity, selected symptoms and physician visits, attitudes toward services received and convenience of obtaining services. In that visits reported were in exclusion of hospital inpatient visits and health department clinic visits, the study may underestimate relative utilization by income groups. Also, because information was collected with reference to the two weeks preceding the interview, one might expect seasonal bias when the results are quoted in annual figures. Some of the conclusions were as follows. There was a redistribution of physician visits per person per year away from upper income earners and towards lower income earners. Implicit in the redistribution by income groups was more visits made by less educated, more visits made by ethnic minorities, and more visits by unskilled workers, the most extreme case being the latter in which unskilled workers showed a 68 percent increase, while executives and professionals revealed a 28 percent decrease in utilization.

1. P. Enterline, "Distribution of Medical Services...", op. cit.

To point out the justification with which increased use was made by the poor, the respondents were asked to report "important symptoms"¹ that were treated by a physician. In addition, the highest income group reported slightly fewer while all other groups reported just about the same number of important symptoms after medicare was made available. But before it was made available the range of symptoms reported was from 59 percent by low income earners to 70 percent by high income earners. With regard to the availability of care, the approximation used was waiting time both for appointment and also once at the physician's office. The results show that low income earners waited fewer days for appointments and less time in doctors' offices after the plan came into effect. The high income earners waited longer for appointments and longer in doctors' offices after the plan started. Moreover, these changes were reflected in questions about the quality of care received. A surprising result here was that 29.7 percent of consumers felt services had worsened as compared to 8.0 percent who felt it had improved. The low income earners were more satisfied than the high income earners, probably for the reasons noted above, 1) increased waiting time in office, 2) increased waiting time for an appointment. When questioned about the quality of care received at the last visit, 91.4 percent of all respondents felt it was the best possible, while the distribution between low and high incomes showed high income earners slightly more satisfied with the quality of care. Also, the disparity between before and after opinions of quality was reduced over income groups.

1. The symptoms judged important were those strongly indicating potential illness or disability.

Thus in terms of the Enterline study of Montreal health services, the impact of a universal system appears to be largely beneficial. In a survey made by the same group of physicians' feelings about the plan it appears that a) no clear quality change occurred, b) physicians were able to spend more time on home visits and streamline office procedures indicating greater quality of care for acute cases; i.e. bed-ridden cases, c) that 1.9 percent rather than 1.1 percent of patients visited for unnecessary calls while 1.8 percent as opposed to 2.6 percent of contacts were made later than they should have been. The redistributive consequences of the plan appear to have been realized without too much damage in other areas of concern.

On the other hand, we might add a number of more searching questions. 1) Is the number of visits a viable and accurate measure of quality or quantity of care? 2) What was the effect on health status of the redistributive effort; this remains uninvestigated? 3) In view of the fact that the low income earners made, on average, three more visits per year than the high income earners, was the redistribution more than projected? 4) The age group under 17 had reduced physician visits by 6.7 percent among low income earning families and 14.8 percent by high income earning families, while usage stayed the same for middle age groups and was reduced for high income earners over 65 by 9 percent and increased for low income earners over 65 by 13.3 percent. The redistribution of service appears to be from the young of all income levels and from the aged of high income earners to the aged of low income earners. Not only does this seem contrary to the intention of the plan but it has obvious income distributional implications. 5) Even though the direction of change was

that desired, is it possible that the income distributional impact of the current program remains regressive due to the higher percentage of income paid by the low income earners for health insurance premiums and taxes?

R.G. Beck

The second relevant Canadian study is that done by R.G. Beck of family health insurance registrations and medical services records from 1963-68 in Saskatchewan.¹ A random sample of 40,000 families was selected from registration files for each year studied and each sample was checked for representativeness by comparison with population distributions of family size. The medical service records contain a patient history file of all physician services that a given patient received under the insurance plan. While registration files revealed family size, age, marital status and address, these data are questionable because pre-1965 a patient identification system was not fully operative and in 1968 a co-payment system was initiated that may have introduced a systematic error. Income of each family was established through a name and address search program of income tax files. Using a strict definition of physician utilization and availability, Beck hypothesized that the degree of non-use of physician services over a year by any income class was inversely proportional to the degree of access that the class enjoyed. He checked both cross-sectional and longitudinal data to reveal the effect of physician services insurance on disparity of accessibility by income group.

1. R.G. Beck, "Economic Class and ...", op. cit.

His results showed: 1) that in the first year of operation of the insurance plan, considerable disparity between income class accessibility existed, 47 percent non-use in lowest income class and 10 percent non-use for highest income class; 2) that the absence of financial barriers increased accessibility more among low income groups than high income groups and that disparity of accessibility was, therefore, reduced over the six years studied, 3) that after six years the degree of non-use appears to be leveling off in such a way that low income groups still display more non-use than high income groups, 30 percent and 10 percent respectively, 4) that both patient elective and physician elective services manifest disparity of accessibility by income class, but over the period studied the patient elective services display relative reductions in non-use levels, while physician elective services show little change.

This study should be regarded with some reservations given difficulties described by Beck in accurate data collection. Also, the definition used by Beck for "accessibility" is questionable in that it ignores availability of services, it gives no indication of need or justification of use, and it fails to distinguish between other socioeconomic characteristics and income. Of major importance for our purposes is the influence that age and family size may have had, age in so far as it is closely related to health status and family size in so far as physician visits are far more likely for large families. It is not clear whether Beck did or did not control for these variables.

The advantage of a "non-use" statistical comparison is that

first visits are normally on patient initiative and this is just exactly what we want to observe; i.e. do certain groups fail to utilize even though financial barriers are absent? Of course, the degree of utilization and the quality are important factors to be determined in greater detail by subsequent studies.

The results above strongly support the hypothesis that financial barriers are of some consequence while raising suspicions of regressive income distribution given the relatively minor changes in disparity between income classes. Beck further suggested grave questions regarding quality of care in that disparity of specialist services and laboratory testing remained large over the period of study and major surgery disparity of non-use actually increase. Note that this result may be simply a reflection of the disparity of physician visits. Above all, the fact that any disparity of use exists over income class raises the question, "Why?" Clearly the objective of the plan is not being realized as services remain inequitably distributed, even though financial barriers are absent. Perhaps a more direct approach to the provision of services for the medically indigent should be investigated?

W.D. Eckstrand

The third Canadian study we will consider is that done by Dr. W.D. Eckstrand of the Manitoba health services using all hospital and and medical fee-for-service data for the year 1970.¹ By dividing the

1. W.D. Eckstrand, "Patterns of Medical and Hospital Care in Manitoba 1970", White Paper on Health Policy, Appendix to Vol. II, (November 1972).

Province of Manitoba into seven regions based on a regional breakdown used for a good deal of statistical analysis regarding development of local economies and population characteristics, Eckstrand was able to correlate broad utilization patterns with health care facility and inventories.

This strong statistical basis for analysis of health services will allow Dr. Eckstrand and Manitoba to fully investigate many facets of health care virtually untouched by researchers in Canada. Future studies planned for Manitoba will relate morbidity patterns to utilization rates and provide rationale for medical manpower planning and facility distribution. It is this kind of data base that is necessary for a complete inquiry into income distribution questions. Already he has been able to analyze regional availability and utilization by studying physician and hospital bed availability by region, interregional flows of patients and intraregional expenditure per capita.

His conclusions were 1) that even after regional utilization was adjusted for age and sex distributions, there was considerable disparity caused by a) ethnic origin (high usage by Indians), b) availability of service -- the surprising result with respect to availability was that "utilization rates are significantly higher for those residents closest to health care facilities"; 2) that much of the disparity in the system was due to physician location preferences and political community pressures rather than regional need criteria; 3) that quality of care as measured by specialist availability was vastly unequitable with over 500 specialists located in Metro Winnipeg and less than 40 specialists in all of the six other regions combined; 4) that hospital bed utilization was far higher in all areas other than Winnipeg and Brandon with admissions per thousand

at 202 for all Manitoba excluding Winnipeg and 162 for Winnipeg.

The main emphasis of Eckstrand's work is that regional disparities in availability of service can have major consequences on both the quality of care given citizens of the regions and the rate at which facilities are utilized. For our purposes his most pressing conclusion was that where certain facilities; i.e. practitioners and specialists, are not available, there may be a good deal of substitution of care procedure significantly affecting the cost of care. The result being that in urban centres where economies of scale can be realized, residents receive far better service as compared to rural districts where care is expensive and residents receive services of reduced quality less often. Although it would be highly coincidental, there does seem to be a system operating that equates cost and benefits from the provider point of view and perpetuates inequities from the consumer point of view.

A.H. Sohn

Finally, a study currently being carried out by Dr. A.H. Sohn will provide a comprehensive investigation of community utilization of hospital services in Toronto.¹ Using an address conversion program, Sohn will be able to correlate O.H.I.P. admission data with census tract data in developing answers to utilization questions. He also will develop a complete survey of hospital facilities available in each community,

1. A. Sohn, "An Analysis of Community Utilization of Hospital Services of Metropolita Toronto and An Analysis of the Service Areas of Hospitals in Metro Toronto: A Research Proposal", Ontario Ministry of Health, 1974.

including acute care, extended care and outpatient care facilities. Although the intention of this study is to provide analysis of the hospital sector with regard to planning needs and questions, the data that is produced should provide answers to income distribution questions with only minor re-examination and amplification.

There remain three items that will have to be included in the study before a complete income distributional picture can be seen. First the study will have to be expanded to all health care programs provided by the insurance scheme and tax subsidy. The main component lacking presently is consideration of physician availability and utilization. Secondly, a measure of quality of care is required that not only identifies what facilities are available but what intervening variables differentiate the degree of success with which these facilities are used. Implicit in this project are two other areas that will yield conclusive answers. These are the measurement of health status of the communities involved before and after contact with the health care facilities. The objective here being to a) evaluate the degree of actual health status benefit received by individuals, and b) relate these measures to the tax burden and insurance payments specific to these individuals.

As discussed earlier the effects of a copayment mechanism may be very important a) in a redistributing health services b) in limiting demand. A study carried out in Saskatchewan by R.G. Beck examines the effect of a different fee on utilization of physicians services by the poor.¹ Using the same data collection system as described earlier Beck

1. R.G. Beck, "On the Effects of Co-Payment on the Poor," Journal of Human Resources, Vol. 9 No. 1, pp. 129-141.

compiled a sample of poor families utilization records for the years 1963 to 1968. In 1968 a deterrent fee of \$1.50 - \$2.00 was authorized to be charged for physician visits. Ignoring the possibility that poor families would not be charged this fee as it was a discretionary charge Beck's demand estimations reveal a general 18% decline in utilization by poor families following the introduction of the deterrent fee. The decline in usage was greater for patient elective than physician elective visits, as expected. This conclusion should not be considered a strictly distributional result in that other income groups were not investigated although such a large variation indicates that the co-payment scheme is excessively regressive.

This same kind of study might be possible in Ontario by examining the effect of the discontinued billing by doctors for the 10% of charges not covered by the insurance scheme. This 10% charge operated much like a deterrent fee except that it would in fact be called a case of coinsurance in as much as the fee paid by the patient is expressed as a percentage of total fees.

V. A RESEARCH AGENDA

In the foregoing discussion of the nature of the distributive impact of government funded health care, we have on several occasions drawn attention to research work that could with advantage be done in this area. There are some research priorities that should be emphasized. First, the depth and quality of data available to the researcher should receive top priority. This job has been started by the Report of the Ontario Council of Health on Health Statistics (1969) but apparently

greater encouragement is needed to cause implementation of the recommendations of that report. Also, since that report there have been a number of new perspectives on health research that alter the data requirements, not the least of which is the rapid growth of health economics. Therefore, the need for input from an economic point of view in the determination of a data collection process is evident. It is noteworthy that the census tract data and the address search program have significantly reduced the size of this project to one of collecting health specific data.

Secondly, central to the issue of types of data collected in the on-going research of health status measurement. A working index of health status is urgently required for the evaluation of health services. The literature in this field is growing rapidly, and many suggestions are being made and are ready for both empirical testing and use. Dr. G.W. Torrance has made many recent contributions in this field and tested them in Ontario. For purposes of defining income distributional impact of health programs, Dr. Torrance's health status scale is well adapted and deserving of consideration.¹ This is an area in which the economist can be part of what is largely a health science team.

Thirdly, the question of quality of care disparity over regional and socio-economic groups is in need of research. For the present, there is little in the way of tested and reliable methodology for evaluating quality of care. On the other hand, the issue has received a number of suggestions for measurement and a good deal of examination of these measures.

1. See Footnote 1 of page 13 of this paper.

The four studies reviewed above, all gave some indication of where the researcher should look for quality of care disparities. The Beck study clearly showed the trend for patient and doctor elective medical care to favour the higher income groups, while the Eckstrand study indicated that physician location preferences and community pressures were the apparent sources of quality disparity in Manitoba. S.M. Miller, in an American study that has been substantiated many times, showed that mental illness was more prevalent in the lower income classes and that the method of treatment for mental illness was also related to social class. Some concepts that should be researched in Ontario are: 1) What are the viable methods of evaluating quality of care for each kind of health program operating in Ontario? 2) Is the quality of care simply a matter of availability distinctions, or is it true that poor quality services exist in lower class districts? 3) How important are quality of care disparities in affecting the benefit and cost of health services? 4) What are and how great is the impact of intervening variables relating socio-economic class and health service effectiveness.

The questions of utilization and availability have been well defined in the literature and by studies already carried out in Canada. Because the results of these studies have been contradictory, there is a clear need for precise and comprehensive examination of these questions in Ontario. The hypothesis that fewer facilities are available in lower classes and regionally dispersed areas is an obvious one. Also of interest, though, is how regional dispersion and unwelcome environments affect the cost of making services available. It has been accepted in Ontario that economic incentives are required to cause equitable dispersion of physicians

and dentists; also, economies of scale affect the cost of providing service to rural areas. These factors should be reviewed regarding their impact on income distribution. Another availability factor seems to be the technical indivisibility of medical services. This matter should be related to the regional population in determining optimal location of services.

With respect to utilization rates there are a great number of determining factors. In general, for income distributional analysis, we want to know just one set of facts. These are the relationships between who pays for and who uses health services. Above all, then, this matter should be examined by relating taxable incomes to family utilization rates. But also, because premiums are paid for the provision of health insurance, the question arises as to how socio-economic, demographic and regional factors affect the utilization rates given that all families pay the same premiums. The latter investigation will be facilitated by the census tract data especially in the identification of interregional patient flows and the socio-economic characteristics of users.

CHAPTER X

Conclusions

The selection of the topics for detailed discussion as was presented in Chapters II through IX represents a judgement of what are thought to be the principal priorities for research by health care economists. Their selection was based on a subjective evaluation of the list of potential research areas using the following criteria: the presumed comparative advantage of economists and their methodology, the likelihood of obtaining sufficient data to carry out the research, and the relevance of the problem and research of it on intermediate term policy. In this last chapter, we propose to discuss some of the general caveats for potential researchers and to mention a limited number of additional research possibilities that are also of potentially significant worth. We do not propose to summarize the conclusions reached with respect to the eight principal areas of research priorities. These are generally stated at the end of each of the Chapters II through IX.

I. General Caveats:

The problem that has probably placed the greatest bottleneck in the way of research on the efficiency of the health care delivery system is the availability of satisfactory data. The problem stems both from the nature of the data that has been collected including its accuracy, comprehensiveness and the activities so described and also from the constraints of confidentiality imposed by the institutions that supply the data. Data describing the resources and activities of the hospital sector are in these respects poor. Similar data describing the non-hospital ambulatory care sector are almost non-existent.

There appears to be a possibility that data on the hospital sector will improve sharply as a result of efforts on the part of Statistics Canada to link data on resources and activities to data on the patients to whom the services

are rendered. The Hospital Medical Records Institute also represents a potentially valuable source of data on the hospital sector. In spite of these expected improvements it seems likely that the researcher must continue to be prepared to allocate a disproportionate amount of his efforts to securing required data. Government must do better in producing data.

As has been noted before, the health care sector is not only a large sector of economic activity but it is also a complex one given the difficulties of defining, let alone measuring, high quality in health care delivery. As a consequence, it seems essential that the researcher educate himself about the peculiarities and constraints of the institutional framework of licensing procedures and accompanying educational requirements and regulations and of the traditional patterns and practices of the economic actors and production units of the sector. Towards this end, the researcher will in all probability have to become conversant with such journals as Medical Care, Inquiry, Health Services Research, International Journal of Health Services, American Journal of Public Health, and so on.

In addition to a degree of immersion in the medical care literature, there may well be significant benefits to the economist in collaborating with members of the more traditional health care research team, especially the epidemiologists, bio statisticians, and public health research groups.

II. Further Research Priorities

In addition to the research priorities discussed previously including those noted in Chapter I, there are a limited number of specific research projects that we might emphasize as deserving of the attention of economists.

A. Denticare:

The prospect of government's entry into the funding of dental care

leads me to think that an evaluation of the impact of such an action on the dental health care delivery system would be of not insignificant value in the very near future. From the standpoint of allocational efficiency, we require an evaluation of the responsiveness of dentists to the increased demand for their services thereby caused. For example, if dentists are by and large working at capacity, the short-run impact might yield relatively small increases in the supply of dental services while at the same time it provided an environment highly conducive to sharply increasing fees.

From the standpoint of distributional equity, we require information on the factors influencing quantity demanded other than the price of the services. If such factors are significant, the removal of the price barrier might lead to even further differences in the consumption of dental services by the highly educated, well to do in comparison to the poorly educated, less well to do.

The discussion in Chapter IX of the distributional impact of government funded hospital care and medical care would likely provide a starting point for such a study of the impact of government funded dental care. We might also draw attention to the several studies of the dental care delivery system that were prepared for the Royal (Hall) Commission and Health Services, 1964 and the Committee on the Healing Arts, 1970.

B. Pharmicare:

As in the case of denticare, the possibility of government funding of a large scale programme to cover the cost of prescriptions is significantly high enough to warrant study of its likely allocational and distributional impact in the health care delivery system. The questions raised with respect to denticare are also relevant in the case of pharmicare. However, the knowledge base from which one starts is perhaps not as well developed. Though the study

of the pharmaceutical industry in the United States has occupied the attention of economists for some time*, the same cannot be said with respect to Canada.** Accordingly, research on the expected impact of a pharmacare scheme probably would have to involve a fundamental study of the pharmaceutical industry.***

C. PARCOST, QUAD and the Drug Benefit Formulary

Of perhaps greater importance than the preceeding research proposal is one that also involves the pharmaceutical industry. Seemingly following the path of the United States in its regulation of new drugs, the Province has established two programmes, first PARCOST (Prescriptions at Reasonable Cost) and subsequently the Drug Benefit Formulary. At the same time, the federal government has embarked on its programme QUAD (Quality assessment programme). The principal objective of these programmes is to provide society with greater protection against unsafe, therapeutically ineffective, and, or, overpriced drugs. At the same time, an apparent unintentional outcome of programmes such as these is the sharp reduction in the rate at which new pharmaceutical preparations are introduced into the health care delivery system.

*For example, see the industry study by H.D. Walker, Market Power and Price Levels in the Ethical Drug Industry, (Bloomington: Indiana University Press, 1971), and the bibliography contained therein; and Regulating New Drugs, edited by R.L. Landau (Chicago: University of Chicago Press, 1973.)

**Three studies have been completed however. See Canada, Department of Justice, Restrictive Trade Practice Commission, Report Concerning the Manufacture, Distribution and Sale of Drugs, (Ottawa: Queen's Printer 1963); Canada, House of Commons, Report of the Special Committee on Drug Costs and Prices (Ottawa: Queen's Printer, 1967); and The Canadian Pharmaceutical Association, Pharmacy in a New Age: Report of the Commission on Pharmaceutical Services. (Toronto: Imperial Press Ltd. 1971).

***In this direction, B. Pazderka is now completing his Ph.D. thesis on advertizing in the pharmaceutical industry for the Department of Economics, Queen's University.

Whether the benefits of these programmes exceed the costs of establishing them including the foregone benefits of new drugs not introduced into the system is an open question.* An evaluation of these programmes would seem to be especially worthwhile.

D. Health Care Needs and the Changing Age Distribution of the Population

The post World War II baby boom appears to have successively led to the building of disproportionately large numbers of new elementary schools, new secondary schools, new universities, new apartment buildings, and so on. At the present time, it is probably a principal cause of the fairly high demands being placed in the housing market. It has presumably had and will have significant impact on the health care delivery system. There would seem to be ample justification for evaluating the nature and magnitude of the likely future impact of changing age distribution on the health care sector. Especially important would seem to be an analysis of the likely demands for health care of the baby boom cohort as it reaches and enters retirement.

E. Support Services

A number of economic activities in the health care sector provide support services, or in the economists' jargon, intermediate products. Ambulance services, laboratory services, radiological services, and so on are examples of these services. With respect to each subset of them, an industry type study would seem to be appropriate in order to provide the basis on which decisions can be made on how to best integrate the provision of these services

*In his study of U.S. regulations, S. Peltzman has found that U.S. society has lost significantly more than it has gained. See S. Peltzman, "The Benefits and Costs of New Drugs", in Regulating New Drugs..., pp. 114-211.

with the provision of the bulk of health care. In particular, what general system of resource allocation, the market price system or government ownership and control for example, is appropriate if the bulk of care is provided with government funding.

III. The Need for Health Economics Research

The health sector in Ontario now accounts annually for resources valued in excess of three billion dollars and amounting to more than six per cent of the gross provincial product. At the same time, the general system of resource allocation now involves government in a major role whereas the market price system has been relegated to a very minor one. The solution to the fundamental economic problem of what, how and for whom must nevertheless continue to be sought; health care resources are scarce, the wants of individuals for health care are insatiable and there are alternative ways of combining health care resources to provide health care. There can thus be little question about the significance of the role of research in health care economics. That research of this kind is so small in relation to the size of the sector is difficult to justify.*

It is worth emphasizing that the economic problem is not one that is solved once and for all; it is a continuing problem involving decisions on the allocational and distributional nature of the health care delivery system. Thus, research on "old" problems may well be a visible and essential element of any new set of priorities for economists in the health care field. Of course, if these "old" problems have been analyzed by researchers from

*Estimates of the negligible amount of research on ways of improving the allocational and distributional efficiency of the health sector are given in R.D. Fraser, "The Economics of Health Research", in Report of Committee on Health Research (Toronto: Ontario Council of Health, 1973). pp. 26-202.

other disciplines with other objectives and methodologies, there might well be further reason for the inclusion of such "old" problems in a new agenda.

